PART I: TOTAL SYNTHESIS OF ASPEVERIN AND PENICIMUTAMIDE A

PART II: TOTAL CHEMICAL SYNTHESIS OF ALL-L AND ALL-D KRAS(G12V) AND THE FURTHER EXPLORATION OF ISONITRILE-MEDIATED PEPTIDE LIGATIONS

A Dissertation

Presented to the Faculty of the Weill Cornell

Graduate School of Medical Sciences

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Adam M. Levinson

January 2017

PART I: TOTAL SYNTHESIS OF ASPEVERIN AND PENICIMUTAMIDE A

PART II: TOTAL CHEMICAL SYNTHESIS AND FOLDING OF ALL-L AND ALL-D KRAS(G12V) AND THE FURTHER EXPLORATION OF ISONITRILE-MEDIATED PEPTIDE LIGATIONS

Adam M. Levinson

Cornell University 2016

Part I:

Fungi serve as a rich source of prenylated indole alkaloids, which exhibit important biological activities including antiproliferative, antibiotic, and antihelminthic properties. Their promise as therapeutics, coupled with their diverse and complex molecular architectures, have made prenylated indole alkaloids popular targets for synthetic chemists in order to probe their activities and develop new synthetic methods. Herein, we describe the first total synthesis of aspeverin, a unique bridged carbamate-containing prenylated indole alkaloid isolated from *Aspergillus versicolor*. We also describe the synthesis of a closely related congener, penicimutamide A, isolated from a mutant strain of *Penicillium purpurogenum*. These molecules belong to a recently described subclass of prenylated indoles thought to be degradation products of parent bicyclo[2.2.2]diazaoctane congeners.

In this research, we showcase a highly diastereoselective Diels−Alder cycloaddition, followed by an electrophilic Rawal arylation – reductive indolization to forge the pentacyclic scaffold of these natural products. A novel sequence for installation of a geminal dimethyl group was also developed. This involved a carefully devised transannular carbamate cyclization followed by a ring-opening /

alkylation sequence. Finally, the bridging carbamate common to both of these natural products was forged using a hypervalent iodine(III)-mediated oxidative cyclization.

Part II:

The KRas protein is a small GTP-binding protein important in cellular signaling, involved in growth, differentiation, motility, and survival. Oncogenic mutated Ras is involved in approximately 30% of all cancers. Despite the identification of this target in human cancers over 30 years ago, no small molecules targeting KRas have been successful in the clinic, and this target has long been considered "undruggable." In an effort to develop novel therapeutics, this project seeks to complete a total synthesis of an all-D amino acid variant of the KRas protein. This synthetic protein enantiomer will be used as a tool for mirror-image yeast surface display studies to identify all-D residue peptide ligands for KRas. In this section, the synthesis of biotinylated variants of KRas(G12V) consisting of all-L and all-D amino acid residues is described. Moreover, we demonstrate that this enantiomeric pair of 166-residue proteins bind to nucleotide substrates as well as Ras-binding peptides with enantiospecificity.

During our studies on the total synthesis of KRas, we further demonstrate the utility of isonitrile-mediated activation of C-terminal thiocarboxylic acids for peptide ligation. This method is advantageous, in that it does not rely on a cysteine disconnection and takes place under mild conditions in polar aprotic organic solvents. Beyond a synthesis of KRas, we also demonstrate that isonitrile-mediated

ligation and native chemical ligation strategies can be combined and applied toward a novel synthesis of HIV-1 protease I66A.

BIOGRAPHICAL SKETCH

Adam Marc Levinson was born in New York, New York in 1988. He spent a majority of his childhood growing up in Atlanta, Georgia. He entered Emory University in 2007, initially on a pre-medical school track. He ultimately changed his major and career interests to chemistry after taking sophomore organic chemistry with Professors Frank McDonald and Simon Blakey. During Summer 2009, Adam was exposed chemical research while working in the neuropharmacology laboratory of Professor Alvin V. Terry, Jr. at the Medical College of Georgia. Gaining a cursory understanding of spectrophotometric enzyme assays with small molecules, he became interested in how to actually make the molecules he was testing. He spent the remaining two years of college working in the synthetic organic chemistry laboratory of Professor Huw M. L. Davies studying asymmetric reactions of rhodium carbenoids. During Summer 2010, Adam was also fortunate enough to work as an intern in process chemistry at Amgen in Thousand Oaks, CA under the direction of Filisaty Vounatsos. His experiences there helped to solidify his career aspirations of discovering new therapeutics to treat disease. After graduating from Emory University with highest honors in 2011, Adam went on to complete as Masters degree at Columbia University, working under the direction of Professor Scott Snyder in the areas of polyphenol and halogenated sesquiterpene natural products synthesis. Adam later transferred in 2013 to the Tri-Institutional PhD Program in Chemical Biology, studying organic synthesis in the laboratory of Professor Samuel J. Danishefsky. Starting in Fall 2016, Adam will be working in Discovery Chemistry at Eli Lilly and Company in Indianapolis, IN.

ACKNOWLEDGMENTS

There are many people to thank for guiding me through my experiences at Memorial Sloan Kettering. Foremost, it has been an honor to learn under the direction of Sam Danishefsky. Having taken his organic synthesis class in 2012, Sam's unique style of teaching helped me to learn not only the basics of named reactions, but more importantly the *strategies and logic* of synthesis. His historical and encyclopedic knowledge of chemistry is truly remarkable, and it has been a privilege to continue learning from him during my time in TPCB. Sam's approach toward research, with a constant eye on changing the lives of patients here at MSKCC through basic science, has taught me to always keep the overarching goals of research in mind. You are an inspirational scientist and person, and I cannot thank you enough for the experience you have given me. I would especially also like to thank Sarah Danishefsky for her encouragement and help during the past few years.

My committee members Minkui Luo, Sean Brady, and Derek Tan have been instrumental to my graduate success. Coming from a more traditional chemistry background, my interactions with them during meetings have challenged me to expand my knowledge and expertise within the field of chemical biology. I thank Gabriela Chiosis for her role in my thesis defense as well. Derek, in particular, has been an incredible mentor to me, and I cannot express enough thanks. Spending time in your chemical biology course, in journal club, and during joint group meetings, you have taught me how to think critically about problems in organic

synthesis and in chemical biology, and you have helped me to become a much better communicator of science.

I would like to give thanks to certain labmates, with whom I have spent a great deal of time learning with. Adam Trotta has been by my side all the way through graduate school at Columbia University and in TPCB. You are an extremely bright and talented guy, and you have taught me a ton of chemistry over the years. I know you will go on to do great things after your PhD. Dr. Steven Townsend was also instrumental to my success when I first joined the lab. Your mentorship helped me to complete my total synthesis project, and I especially enjoyed talking about and eating Pizza Park with you! Working with Dr. Andrew Roberts has also been invaluable to my education. Thank you for teaching me the basics of peptide chemistry, for helping me to choose the right experiments, for providing endless banter on synthesis strategies and reaction mechanisms, and for pushing on me your perfectionist attitude when it comes to papers and presentations. Thank you also for selflessly and painstakingly reading this entire thesis multiple times. You are a great friend, and you have made me a much better chemist than I would otherwise be. I look forward to seeing the science that comes out of your independent career! To the rest of the current Danishefsky lab – Dr. Baptiste Aussedat, Dr. Bill Walkowicz, and Dr. Abram Axelrod – thank you for providing a great work environment. You have made graduate school a fun place to be, and I know you will all be very successful.

For the second part of my described work, I must acknowledge the other Danishefsky lab members who have taken part in building this project. This

includes Dr. Gardner Creech, Dr. Ting Wang, Dr. Andrew Roberts, and Dr. Michael Peterson. Dr. John McGee was instrumental to our project, contributing biochemical and biophysical characterization of our synthetic proteins. I thank him and Professor Gregory Verdine for kindly devoting their time and energy to this project.

I need to thank my family, Jane, Keith, Josh, and Amy for raising me to achieve the best I can and for making me who I am. I would not be at this point in school without your support, love and encouragement throughout the years. I would also like to thank my in-laws, Sally, Henry, Bob, and Marge. It has been great being in the same city with you for all these years, and I am so lucky to have four more parents that I can truly call parents. To my dog, Dooley- Thank you for filling my years of graduate school with endless amounts of amusement, and for your constant reminder not to take life too seriously! Finally, I need to thank my wife, Dorothy Abrams. You have been my rock and guide throughout graduate school. You know as well as I do, that your support throughout this process has brought me to where I am. You are my best friend, my voice of reason, and the best partner I could ask for in life. I can't wait to see what our future years in Indianapolis hold for us!

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

LIST OF SCHEMES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

LIST OF TABLES

Chapter 1

Chapter 3

LIST OF ABBREVIATIONS

Acm	acetamidomethyl
Ac	acetyl
AcOH	acetic acid
AcCl	acetyl chloride
AIBN	azobisisobutryonitrile
AIDS	acquired immunodeficiency syndrome
Alloc	allyloxycarboneyl
Bn	benzyl
BnBr	benzyl bromide
Boc	*tert*-butyloxycarbonyl
COSMIC	Catalog of Somatic Mutations in Cancer
CNS	central nervous system
Da	Dalton
DBU	1,8-diazabicycloundec-7-ene
DCC	*N,N'*-dicyclohexylcarbodiimide
DIBAL-H	diisobutylaluminum hydride
DIC	*N,N'*-diispropylcarbodiimide
DIEA	diisopropylethylamine
DKP	2,5-diketopiperazine
DMA	*N,N*-dimethylacetamide
DMAPP	dimethylallyl pyrophosphate
DMAP	4-dimethylaminopyridine
Dmb	2,4-dimethoxybenzyl
DMF	*N,N*-dimethylformamide
DMSO	dimethylsulfoxide
DOS	diversity-oriented synthesis
DPPA	diphenylphosphoryl azide
dppf	1,1'-Bis(diphenylphosphino)ferrocene
dr	diastereomeric ratio
DTS	diverted total synthesis
DTT	dithiothreitol
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDT	1,2-ethanedithiol
EGFR	epidermal growth factor receptor
ESI-MS	electrospray ionization-mass spectrometry
Et	ethyl
FCMA	formimidate carboxylate mixed anhydride
FDA	Food and Drug Administration
Fmoc	9-fluorenylmethyloxy carbonyl
FTI	farnesyl transferase inhibitor
GAP	GTPase activating protein
G-CSF	granulocyte colony stimulating factor
GDP	guanosine diphosphate
GEF	guanine nucleotide exchange factor

GGTI	gerganylgeranyltransferase inhibitor
Gnd·HCl	guanidinium hydrochloride
GppNHp	Guanosine 5′-[β,γ-imido]triphosphate
GTP	guanosine triphosphate
H3R	histamine H3 receptor
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HBTU	2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HBS	hydrogen bond surrogate
HIV	human immunodeficiency virus
HFIP	1,1,1,3,3,3-hexafluoroisopropanol
Hmb	2-hydroxy-4-methoxybenzyl
HOBt	hydroxybenzotriazole
HOOBt	hydroxyl-3,4-dihydro-4-oxo-1,2,3-benzotrazine
HOSu	*N*-hydroxysuccinimide
HPLC	high performance liquid chromatography
HVR	hypervariable region
Hz	Hertz
IMDA	intramolecular Diels−Alder reaction
KAHA	ketoacid-hydroxylamine
KCL	kinetic chemical ligation
KHMDS	potassium bis(trimethylsilyl)amide
LRMS	low resolution mass spectrometry
MDR	multidrug resistance
Mant	2-methylanthraniloyl
Me	methyl
MeCN	acetonitrile
MeOH	methanol
MFD	metal free desulfurization
MPAA	4-mercaptophenyl acetic acid
Mpe	3-methyl-pent-3-yl
NCL	native chemical ligation
NMR	nuclear magnetic resonance
NPIF	*o*-(nitrophenyl)phenyliodonium fluoride
NRPS	non-ribosomal peptide synthetase
OMER	*ortho*-mercaptoaryl ester rearrangement
ORTEP	Oak Ridge Thermal Ellipsoid Plot
Oxyma	ethyl(hydroxyimino)cyanoacetate
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
Ph	phenyl
PhMe	toluene
PI3K	phosphoinositide 3-kinase
PIDA	(diacetoxyiodo)benzene
PIFA	[Bis(trifluoroacetoxy)iodo]benzene
PPI	protein-protein interaction

Pydec	2-pyridyldithioethyloxycarbonyl
*p*TsCl	*para*-toluenesulfonyl chloride
*p*TsOH	*para*-toluenesulfonic acid
RBD	ras-binding domain
RTK	receptor tyrosine kinase
SAR	structure-activity relationship
SEC	size exclusion chromatography
SOS	son of sevenless
SPPS	solid-phase peptide synthesis
TBS	*tert*-butyldimethylsilyl
^{t}Bu	*tert*-butyl
^{t}BuNC	*tert*-butylisocyanide
TCEP	tris(2-carboxyethyl)phosphine
TES	triethylsilyl
TFA	trifluoroacetic acid
TFAA	trifluoroacetic anhydride
TFE	2,2,2-trifluoroethanol
TfOH	trifluoromethanesulfonic acid
THF	tetrahydrofuran
TMSE	trimethylsilylethyl
Trt	trityl
UPLC	ultra performance liquid chromatography
VA-044	2,2'-Azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride
YSD	yeast surface display

Chapter 1: Introduction for the Total Synthesis of Aspeverin and Penicimutamide A

Motivation for the study of Natural Products Total Synthesis

Natural products represent an important class of molecules rich in biological activities and functions that serve as vital inspiration for drug discovery.[1] These secondary metabolites derive from plants, fungi, microbes, and higher-order life forms and often play critical biological or survival roles for the organisms from which they are biosynthesized.[2] Prior to the age of modern medicine, plants and extracts containing these naturally-derived molecules were widely used for their medicinal value, with the earliest accounts of their use dating back to 2600 BC.[3] Today, many natural products are counted among the most valuable and historically utilized therapeutics, with broad representation in the categories of antineoplastic agents, antibiotics, antivirals, antihelminthics, immunomodulators, CNS-modulators, and cardiovascular agents, among others. Notable examples of such molecules are highlighted in **Figure 1**.

Owing to their high density in sp^3 hybridized centers and diverse functional patterns, natural products often display high target specificity and biological functions, and often serve as exceptions to Lipinski's rule of five governing the "druggability" of a given molecule.[4] These anticipated properties inspire the generation of "natural product-like" molecular libraries in drug discovery; a lofty ongoing goal, which is realized to some extent with advances in synthetic strategies such as diversity-oriented synthesis (DOS).[5] Advances in biological strategies such as combinatorial biosynthesis, metagenomic analysis, and synthetic biology have

also aided in the discovery of new "unnatural products" or previously unattainable natural products.[6,7,8] According to a 2016 analysis of all small-molecule drugs approved by the United States FDA between 1981-2014, 65% of these 1211 chemical entities are either natural products themselves, derived synthetically from natural products, or were inspired by natural product pharmacophores.[9]

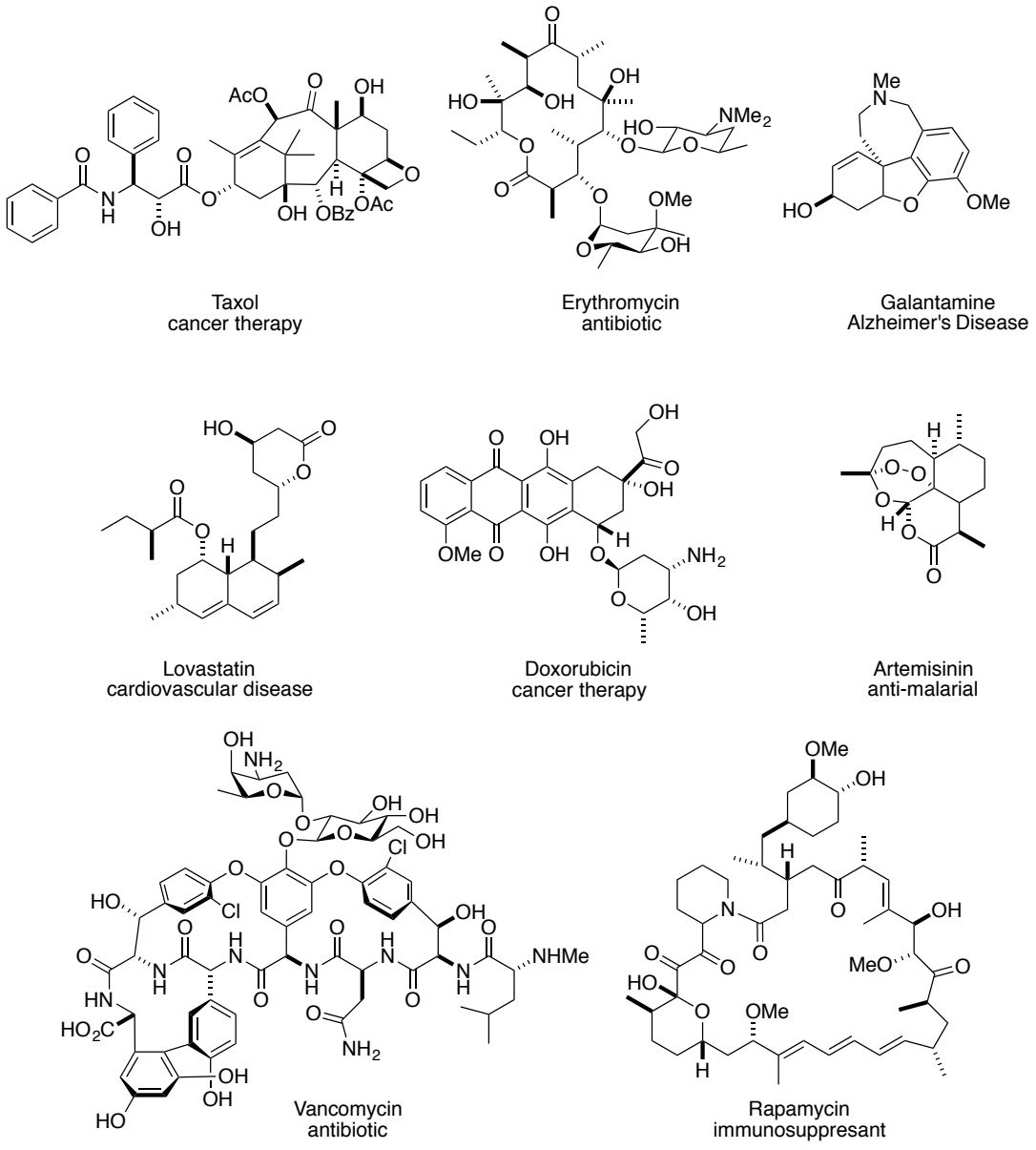


Figure 1: Representative natural products with important medicinal value

Modern organic synthesis has endowed chemists with the tools to achieve total syntheses of such important molecules—many densely functionalized, challenging targets—starting from simpler stock chemicals. Complex target-oriented studies, although narrowly focused, open the door for the exploration of the biology of simplified pharmacophores or diversified analogs with preferable metabolic and toxicological profiles. This process, previously termed by our laboratory as diverted total synthesis (DTS), allows for systematic structure-activity-relationship (SAR) evaluation of complex natural products.[10] In this way, accessible synthetic core structures can be edited, simplified and evaluated at will. The logic of DTS empowers the chemist with full control of every atom within a target of interest – an ultimate goal in drug discovery that cannot be fully realized through reliance on biosynthetic machinery or semisynthetic approaches (**Figure 2**). Shown in **Figure 3** is a notable example benefited by this logic which was devised in our laboratory. Epothilone B (EpoB, **1**) is a highly cytotoxic polyketide macrocyclic natural product, which interferes with tubulin similar to Taxol, and exhibits potent cytotoxicity to multidrug resistant (MDR) cell lines.[11] Our laboratory completed the total syntheses of epothilones A and B in 1996 and 1997, in which EpoB was found to be extremely toxic at low doses in mouse models, precluding its utility as a drug candidate.[12,13,14] However, the development of a *de novo* synthesis from simple building blocks enabled the identification of a reactive epoxide moiety responsible for this toxicity. Synthetic editing of this functionality to a simple olefin, or to the corresponding trifluoromethyl olefin (**3**) significantly reduces this broad toxicity, while maintaining its potent antitumor effects.[15,16]

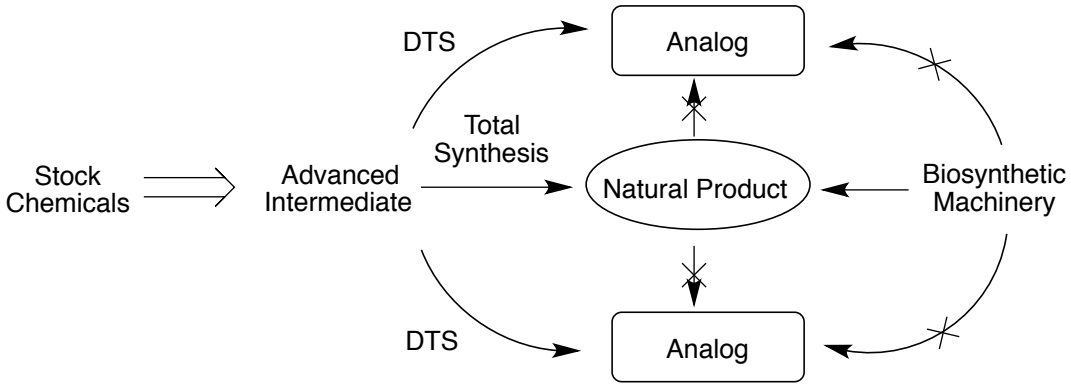


Figure 2: Concept of diverted total synthesis (DTS)

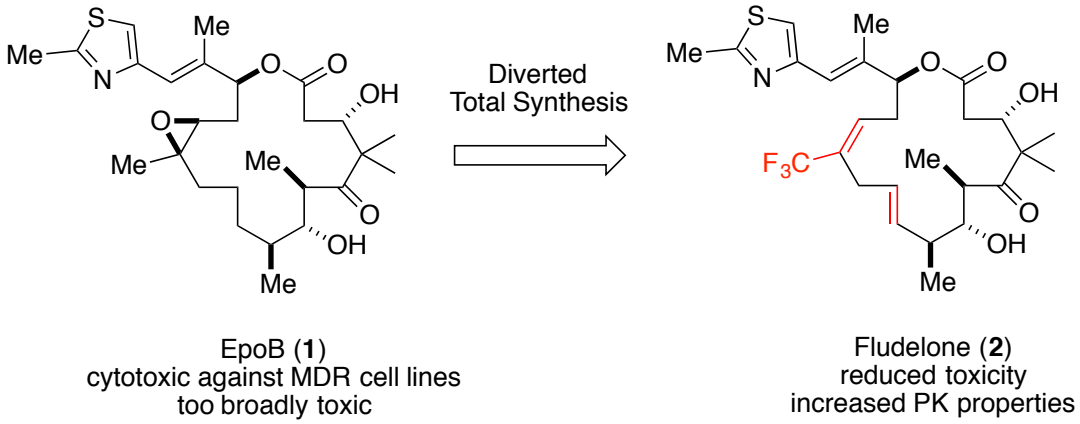


Figure 3: Application of DTS logic to the discovery of fludelone (**2**)

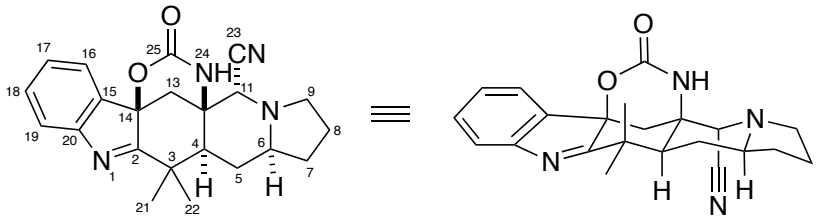
The study of natural products total synthesis bears benefits beyond providing quantities of natural or non-natural analogous compounds for biological study. Due to their complexity, natural products present a formidable synthetic challenge, which challenges the capabilities of modern synthetic organic chemistry.[17] Historically, the challenges associated with synthesizing such molecules have served as inspiration for the development of new, generally-applicable methods in organic chemistry. Arguably, early examples of natural products total synthesis in

steroid and alkaloid chemistry inspired our current understanding of synthesis logic for generating carbon-carbon bonds and our knowledge of heterocyclic chemistry, respectively. For example, research on the targeted synthesis of β-lactam antibiotics inspired the development of ubiquitous amide bond synthesis methods centered on carbodiimide-based reagents – methods which have broader impact beyond these individual targets – especially in the context of peptide and protein chemical synthesis.[18] In the modern age, natural products continue to inspire the development of stereo- and enantioselective synthetic methods, as well as the discovery of new catalytic methods.

Aspeverin (3) : A Prenylated Indole with Unique Synthetic Challenges

Aspeverin (**3**, **Figure 4**) is an indole alkaloid isolated by Ji and co-workers from an endophytic species of *Aspergillus versicolor* found in a *Codium fragile* species of marine green algae in China.[19] After 30 days of fermentation (15L fermentation broth), extraction afforded 29.6 g of ethyl acetate-soluble organic material, from which 8.6 mg (0.029% yield) of pure **3** was ultimately isolated. In particular, the molecule contains distinctive functionality among alkaloids isolated from similar fungal species. The bridging carbamate, linking the indole C-3 carbon (position 14 within the natural product numbering) with an angular amine is particularly unique to this natural product; an oxidative linkage which has not been previously observed. Additionally, the α-cyanoamine functionality is a rarely observed moiety in alkaloid natural products.

The isolation chemists reported preliminary biological data for this molecule, which includes antibiotic activity as well as inhibitory properties of phytoplankton and marine zooplankton growth (**Table 1**). Aspeverin (**3**) displays moderate inhibitory activity against the plankton species *A. salina* and *H. akashiwo*. Of significance, Aspeverin (**3**) also displayed antibiotic activity against Gram-negative species *P. mirabilis, E. cloacae*, as well as Gram-positive *B. cereus*.



aspeverin (**3**)

Figure 4: Structure of aspeverin **3**.

Table 1: Preliminary inhibition of phyto- and zooplankton growth and antibiotic activity displayed by **3** (adapted from Ref. 19)

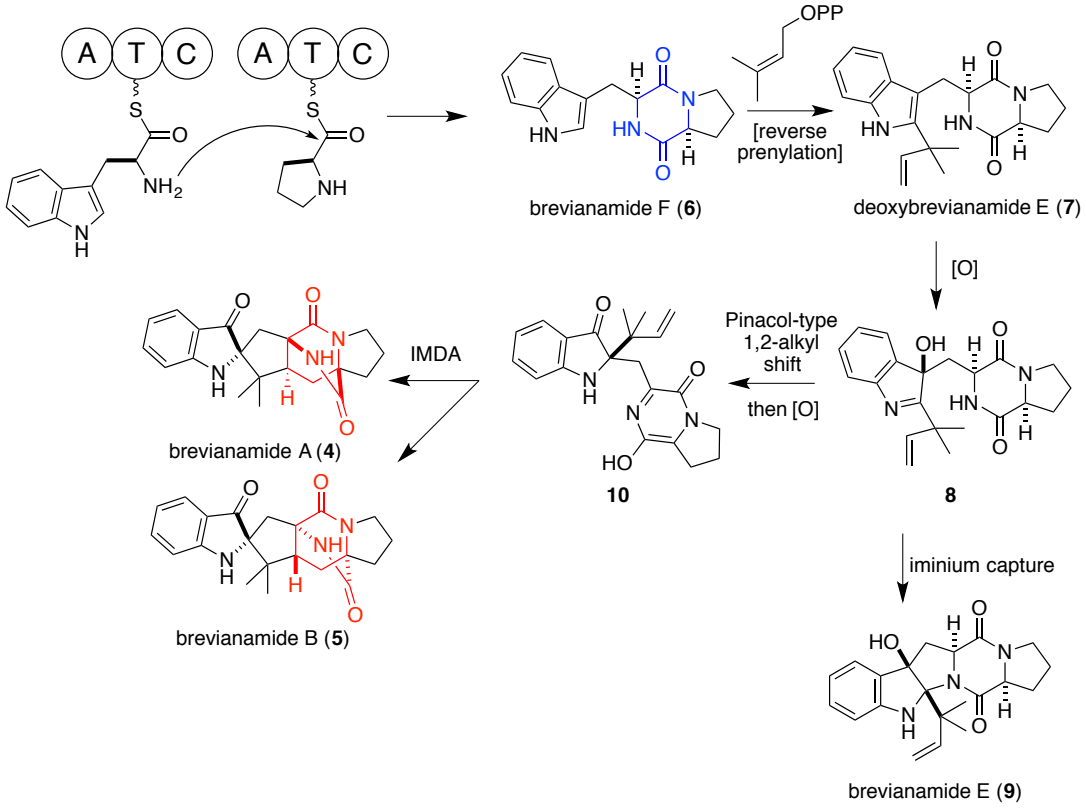
Species	3	$K_2Cr_2O_7$	chloramphenicol
Artemia salina (lethal rate at 100μg/mL; 24 hr)	27.5%	100%	-
Heterosigma akashiwo (EC_{50}, μg/mL; 24 hr)	6.3	13.5	-
Heterosigma akashiwo (EC_{50}, μg/mL; 96 hr)	3.4	5.1	-
Proteus mirabilis (inhibitory diameter, mm; 24hr)[a]	7.0	-	26
Eterobacter cloacae (inhibitory diameter, mm; 24hr)[a]	6.5	-	22
Bacillus cereus (inhibitory diameter, mm; 24hr)[a]	6.0	-	20
Vibrio ichthyoenteri (inhibitory diameter, mm; 24hr)[a]	0	-	28

[a] 30 μg **3**/disk used

The structurally related natural products pertinent to the subject of this thesis are fungal-derived prenylated indole dipeptide alkaloids. *Aspergillus* and *Penicillium* fungi, in particular, serve as a rich source of prenylated indole alkaloids.[20,21,22] Biosynthetically, these dipeptidyl alkaloids are derived from L-tryptophan and one or more amino acids (commonly cyclic amino acids such as

proline or pipecolic acid, although not limited to these) by means of non-ribosomal peptide synthetases (NRPS's). Dipeptide indole alkaloids biosynthesized through this process generally contain a characteristic 2,5-diketopiperazine (DKP) ring formed from cyclization of a free amine onto an aminoacyl thioester intermediate (**Scheme 1**).[23] Due to the modular reactivity of the indole heterocycle within tryptophan, relatively simple common dipeptide intermediates are rapidly diversified through oxidation and prenylation with variable stereo- and regiochemistries, forging progeny with complex and distinct molecular architectures. To highlight one such example, **Scheme 1** outlines the biosynthetic pathway leading to brevianamides A and B (**4** and **5**).[24] Brevianamide F (**6**) serves as a precursor to many natural products, bearing the signature DKP ring (highlighted in blue) formed between L-Trp and L-Pro. Prenylation can occur under a variety of modes with one or multiple units of dimethylallyl pyrophosphate (DMAPP); one example being the conversion of **6** to deoxybrevianamide E (**7**) by means of a reverse prenyl transferase. Oxidation of the indole ring system leads to hydroxyindole **8**, which subsequently can undergo nucleophilic addition to afford brevianamide E (**9**), or alternatively can undergo rearrangement to afford a spiro-indoxyl ring. Upon further oxidation, **10** is poised to undergo an intramolecular Diels-Alder reaction (IMDA), forging the diastereomeric bicyclo[2.2.2]diazaoctane-containing natural products brevianamide A and B (**4** and **5**, ring system highlighted in red). Deoxybrevianamide E (**7**) itself is also capable of undergoing an oxidative Diels-Alder reaction to generate a bicyclo[2.2.2]diazaoctane skeleton without prior indole rearrangement, giving rise to non-rearranged indolic structures such as the

malbranchiamide natural product class (*not shown*).[25] Not only rich in structural diversity, these metabolites exhibit a spectrum of biological activity, including potent antiproliferative, antibiotic, and antihelminthic activity. Such structural complexity and biological activities have historically prompted a great deal of interest from synthetic chemists in this area of natural products, including several successful efforts from our own lab.[26-29] In addition to brevianamides A and B, **Figure 5A** highlights several other example classes of bicyclo[2.2.2]diazaoctanes found in Nature.



Scheme 1: Representative biosynthesis of diverse prenylated indole dipeptide alkaloids of the brevianamide family, derived from L-Trp, L-Pro, and DMAPP.

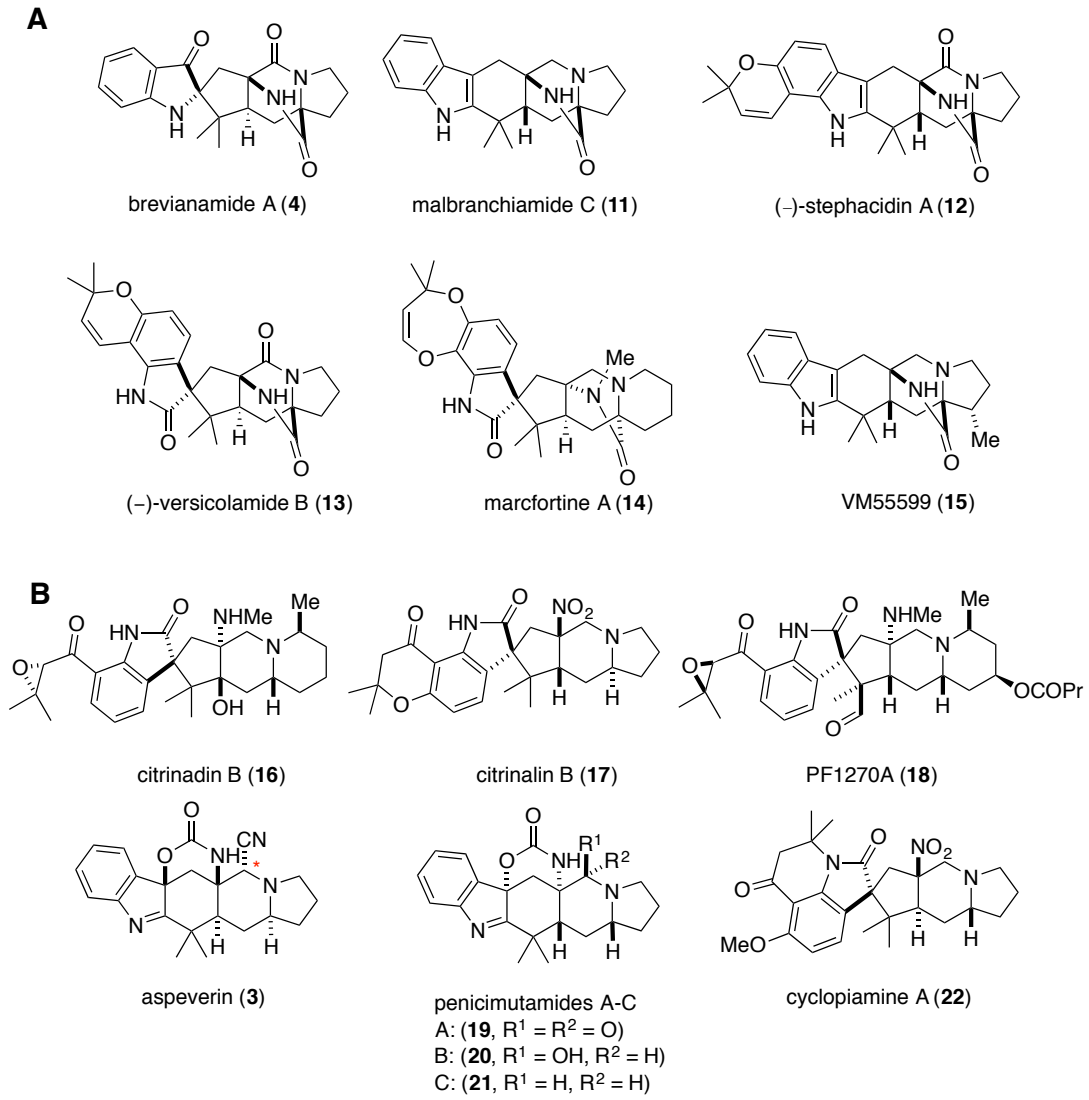


Figure 5: (**A**) Bicyclo[2.2.2]diazaoctane-containing prenylated indole dipeptide natural products. (**B**) Structurally related prenylated indoles lacking the bicyclo[2.2.2]diazaoctane ring system.

In recent years, a number of structurally similar fungal natural products *lacking* this bicyclo[2.2.2]diazaoctane ring system have been identified (**Figure 5B**). These molecules are structurally comprised of the same biosynthetic precursors, but appear to be degradation products of parent bicyclo[2.2.2]diazaoctane congeners. Structurally, aspeverin (**3**) belongs within this category of natural products, although its speculative biosynthesis is not confirmed. Likely components that can be

structurally recognized include Trp, Pro, and one unit of DMAPP. The α-cyanoamine moiety is unique, and may arise from a partial reduction at the highlighted carbon atom (C-11, as outlined in **Figure 4**); a position which is commonly oxidized as the lactam or reduced to a methylene in similar congeners. In fact, during the course of writing this thesis (July 2016), Li and co-workers reported the isolation of penicimutamides A-C (**19-21**) from a mutated strain of *Penicillium purpurogenum*, which differ from **3** only at C-11 with varying oxidation states.[30] These researchers suggest a similar biosynthetic pathway as postulated here, and also suggest that the absolute stereochemistry of aspeverin was likely assigned incorrectly. Such examples of other hypothesized related molecules are highlighted, and include citrinadin B (**16**), citrinalin B (**17**), and PF1270A (**18**). In 2005, Mugishima and co-workers proposed that citrinadin B (**16**) may arise from lactam ring-opening of bicyclo[2.2.2]diazaoctane-containing hypothetical precursor **23**, with loss of a carbonyl in the form of carbon dioxide to obtain the observed ring-opened scaffold observed in the natural product (**Figure 6**).[31] In 2014, a largely collaborative isolation, total synthesis, and biosynthesis study similarly conjectured the existence of citrinalin C (**25**) as a hypothetical bicyclo[2.2.2]diazaoctane-containing precursor to known metabolites citrinalins A and B (**24** and **17**).[32] This hypothetical precursor was ultimately isolated as a minor metabolite from *P. citrinum* F53, lending credence for their assumption, as well as indirect support for the existence of bicyclo[2.2.2]diazaoctane-containing precursors to molecules such as **3** and **16-22**.

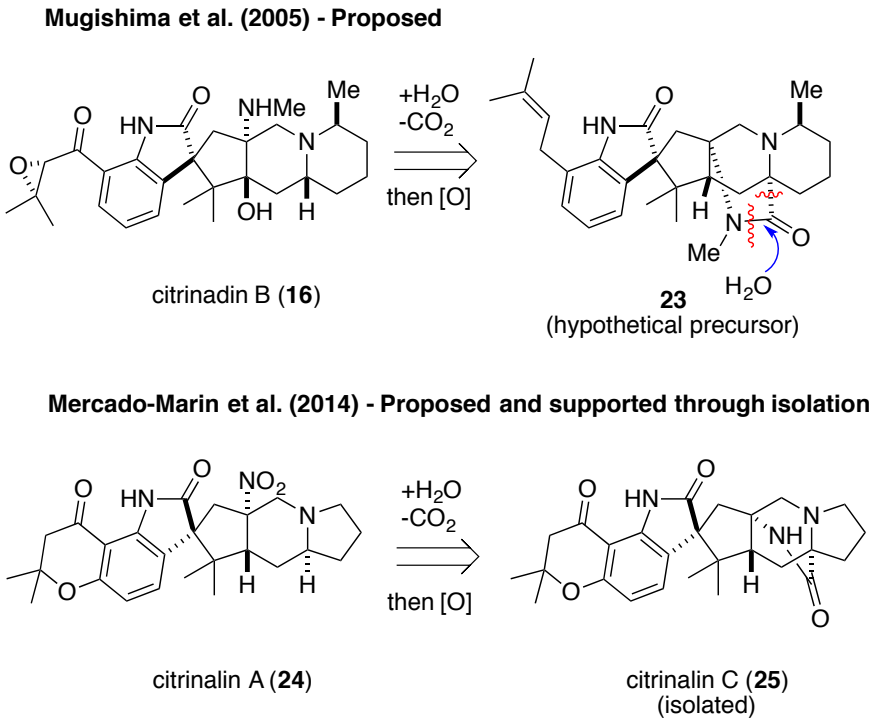
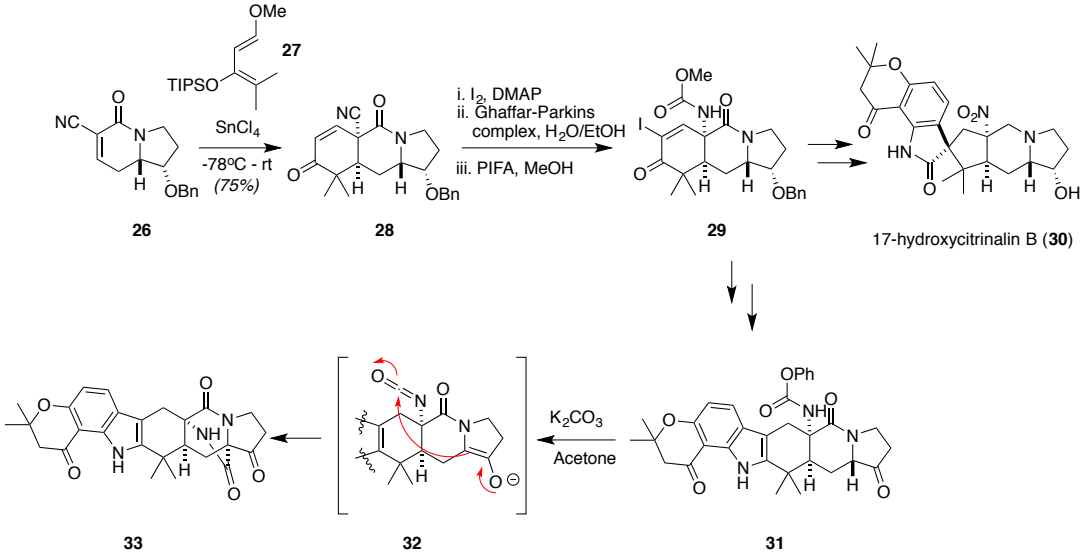


Figure 6: Biosynthetic hypotheses for arisal of bicyclo[2.2.2]diazaoctane-deficient natural products.

Despite myriad total synthesis studies revolved around prenylated indole alkaloids, the highlighted unique ring-opened natural products present novel challenges for stereoselective synthesis strategies. Arguably, to approach the synthesis of these molecules from higher-order bicyclo[2.2.2]diazaoctane precursors would be quite difficult, and previously described methods for the synthesis of DKP-containing indole alkaloids may not be readily applicable to the synthesis of this class of molecules. A number of groups, including efforts by the Wood, Martin, Sarpong, and Sorensen groups, either achieved or aimed at syntheses of citrinadin A (*not shown*) and citrinadin B (**16**).[33-37]

Due to the structural similarities of their targets and differences in synthetic challenges associated with the target of this study (**3**), it is instructive to describe the key elements of Sarpong's unified strategy toward the synthesis of citrinalin family

members as well as their connection with bicyclo[2.2.2]diazaoctane-containing congeners.[38] Highlighted in **Scheme 2** is their general strategy, which begins with a highly diastereoselective Diels-Alder reaction between indolizidine **26** and diene **27** to afford a single diastereomer **28**. Conversion of the angular cyano group to the corresponding amide, followed by oxidative Hoffmann rearrangement mediated by [Bis(trifluoroacetoxy)iodo]benzene (PIFA) affords carbamate **29** as a key intermediate. This intermediate was further elaborated to obtain 17-hydroxycitrinalin B (**30**), bearing an angular nitro group and rearranged spiro-oxindole ring. Alternatively, **29** could be elaborated to phenyl carbamate **31**, which underwent base-catalyzed enolization and isocyanate-capture, ultimately yielding the bicyclo[2.2.2]diazaoctane scaffold related to the natural products (+)-stephacidin A (***ent*-12**) and (+)-notoamide I (*not shown*).



Scheme 2: Sarpong's unified approach toward prenylated indole alkaloids

A close comparative analysis of various related prenylated indole dipeptide alkaloids yields subtle, but notable structural differences between the targets

highlighted in **Scheme 2** and the subject of this study (**3**). Analysis of relative stereochemical relationships in **Figure 7** reveals the distinguishing *syn* relationship at the C:D ring junction within the targets found in Sarpong's work, as well as the *anti* relationship between the distal angular hydrogen atoms (**17** as one example). Aspeverin (**3**) as well as penicimutamides A-C (**19-21**), on the other hand bear an *anti*- C:D ring junction and a *syn* relationship between the highlighted angular hydrogen atoms. A similar relationship is echoed in related natural products such as PF1270A (**18**).[39] This *anti* relationship at the C:D ring junction is quite rare within the bicyclo[2.2.2]diazaoctane natural products, only observed in the case of the brevianamides (**4, 5**), versicolamide B (**13**), and chrysogenamide A (*not shown*). The majority of bicyclo[2.2.2]diazaoctane natural products instead bear a *cis* ring junction, as highlighted for the case of VM55599 (**15**). These differences, especially with respect to the stereochemistry of the C:D ring juncture are critical to the approaches necessary for their total synthesis. For example, the Diels-Alder reaction between **26** and **27** incorporates the distinctive geminal dimethyl group and gives rise to the *cis* C:D ring juncture in a straightforward manner. Moreover, the *trans*-relationship between angular hydrogen atoms sets the stage to attain molecules such as citrinalin B (**17**) and stephacidin A (**12**). This direct approach would be difficult to apply to members with a *trans* C:D ring juncture such as aspeverin (**3**) or PF1270A (**18**).

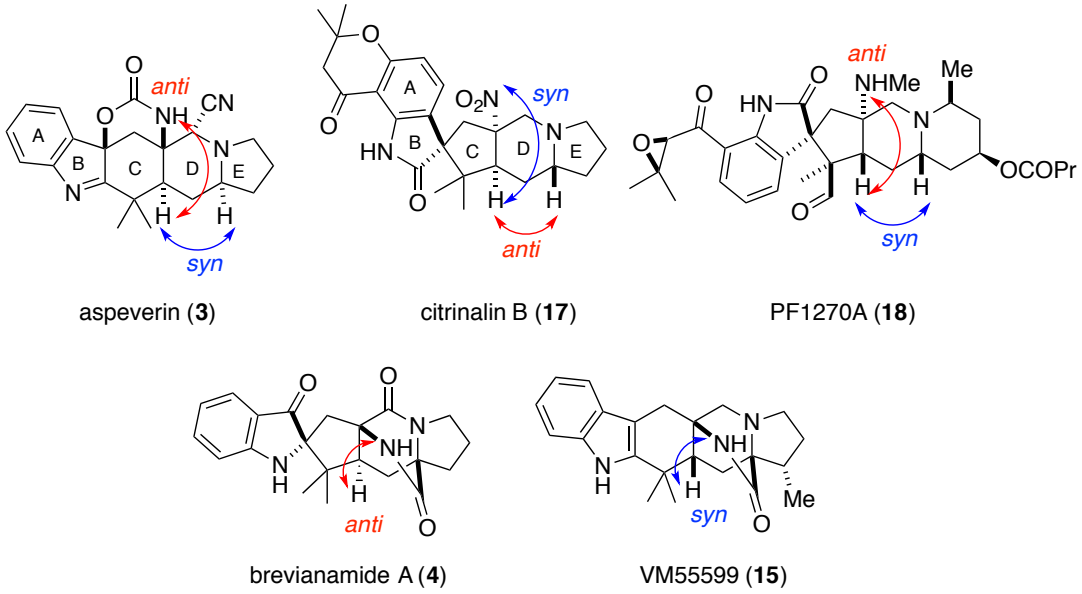


Figure 7: Comparative stereochemical analysis of structurally similar indole alkaloids

Proposed Synthetic Access to Aspeverin (3)

Retrosynthetically, our approach toward the synthesis of aspeverin relied on a late-stage construction of the unique cyclic carbamate linking the indole C3-carbon to the angular nitrogen atom (**Figure 8**). Although one could envision the synthesis of this cyclic carbamate via the linking of a C3-hydroxyindole with a free angular amine using a phosgene-like reagent, such a disconnection could be complicated by a lack of stereocontrol in the indole oxidation, possibly giving rise to diastereomers. Additionally, hydroxyindoles are known to be susceptible to rearrangement, either to form spiroindoxyl or spiro-oxindole products depending on the electronic nature of the adjoined aryl ring. Therefore, our disconnection relied on a less obvious approach, whereby a pendant carbamate (**34**) could be used to directly cyclize onto the indole core structure, negating any potential issues in stereoselectivity or rearrangements. In this plan, the indole reacts as an electrophile,

in an umpolung fashion. This is in contrast to the generally understood nucleophilic character of indoles. This direct intramolecular cyclization of a carbamate onto the C3-position of indole simplifies the target and is supported by literature precedent with regard to the electrophilic reactivity of indoles. However, the exact transformation is unknown, and our proposal served as motivation to develop this new bond-forming reaction.

Further disconnection of **3** by partial lactam reduction and stereoselective cyanation leads back to compound **34**. The proposed cyanation of an intermediate hemiaminal derived from reduction of the parent lactam was predicted to proceed with a high level of stereoselectivity arising from trans-diaxial attack of cyanide onto the iminium intermediate.[40] Installation of the angular carbamate was hypothesized as arising through a stereospecific Curtius-type reaction of the corresponding acyl azide, or alternatively through Hoffmann reaction of the corresponding amide at that position. This simplification brought us to core scaffold **35**. Taking advantage of the angular electron-withdrawing carboxylic acid derivative, **35** was ultimately traced back to **36**, arising from a Diels-Alder reaction between **37** and **38**. This disconnection was advantageous because of the facile entry it would allow toward the core CDE ring system of aspeverin (**3**). However simplifying, a Diels-Alder approach would also create challenges in the synthetic route. Most obviously, the highlighted stereocenter of **36** (highlighted in blue) would require inversion from a *cis* ring junction to the desired *trans* relationship (highlighted in red for **35**). A geminal dimethyl group would require later installation in order to accommodate this stereochemical inversion. Further

discussion of our plans for these necessary transformations will follow later in this chapter.

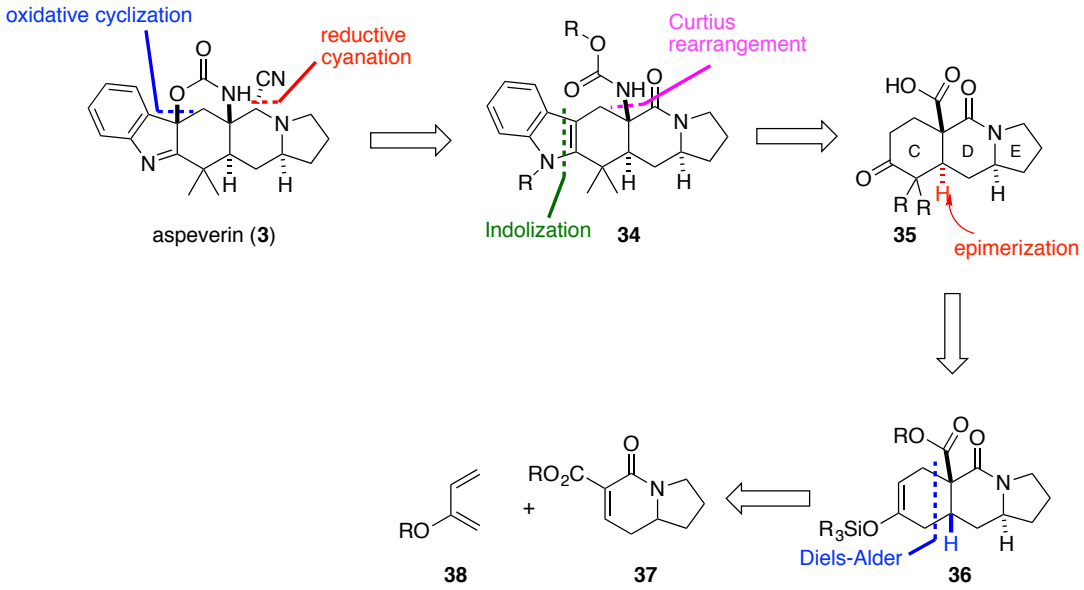


Figure 8: Retrosynthetic disconnections for aspeverin (**3**)

Umpolung Indole Reactivity: Precedent for the Oxidative Carbamate Cyclization

Explicit examples of umpolung indole reactivity are highlighted in **Figures 9** and **10**. One of the most common methods for oxidation of the indole core is through the use of hypervalent iodine(III) reagents such as $PhI(OAc)_2$ (PIDA) or $PhI(OTFA)_2$ (PIFA). In its simplest form, Awang and co-workers reported in 1980 that treatment of 2,3-dimethylindole (**39**) with $PhI(OAc)_2$ in the presence of an alcoholic solvent such as methanol yields 3-methoxy-2,3-dimethylindolenene (**41**) in 35% yield.[41] The authors postulated that the oxidative nucleophilic capture takes place via intermediate **40**, wherein the indole nitrogen atom becomes a ligand for the hypervalent iodine reagent. More complex instances of this reactivity have been

demonstrated since this initial discovery. Treatment of indole **42** with $PhI(OAc)_2$ initiates intramolecular cyclization of a tethered dihydrooxazole to ultimately give rise to a pair of diastereomeric bis-annulation products **44** and **45**.[42] Substoichiometric amounts of oxidant in a non-nucleophilic solvent can produce indole dimerization products in a variety of modes, whereby half of the indolic material undergoes oxidation, and a second equivalent of indole acts as an intermolecular nucleophile. One such example is the dimerization of 1,2,3,4-tetrahydrocarbazole (**46**) initiated by $PhI(OAc)_2$ (0.5 equiv.) to produce indolenine dimer **48** as an equal mixture of diastereomers.[43]

Although mechanistically dissimilar from the previously described examples, Feldman and co-workers have extensively investigated the vinylogous Pummerer-like reactivity of 2-thiophenylindoles to incorporate nucleophiles at the indolic 3-position.[44] This process is initiated either through traditional Pummerer-rearrangement conditions (*i.e.* activation of a sulfoxide with TFAA or Tf_2O) or more efficiently through direct engagement of the sulfide with hypervalent iodine(III) reagents. Compound **49** reacts with Stang's reagent, PhI(CN)OTf, to promote intermolecular cyclization of a tethered *tert*-butyloxycarbamate, yielding **51** in 74% yield.[44,45]

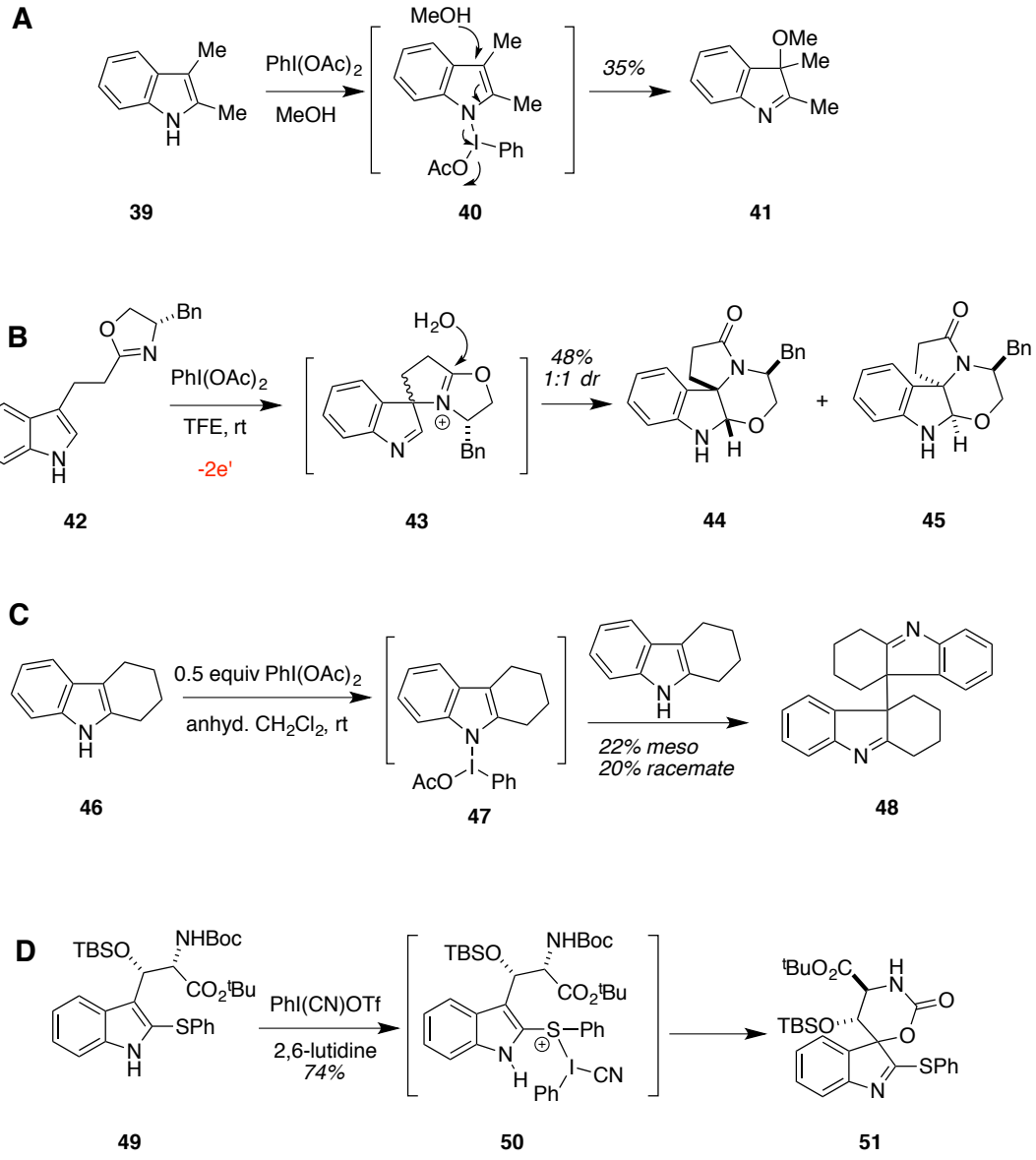


Figure 9: Hypervalent-iodine mediated oxidation of indoles to promote nucleophilic capture at the C3 position.

Alternative to oxidative nucleophilic capture reactions using hypervalent iodine(III) reagents, the C-3 position of indole can also be functionalized via pre-oxidized intermediates. Among others, this can take the form of 3-haloindolenine intermediates or *N*-hydroxyindoles. Upon activation via acetylation or sulfonylation, Somei and co-workers have reported that *N*-hydroxyindoles can undergo N-O bond

fragmentation, to yield 3-functionalized indolenines.[46] This is observed in the C-3 acetoxylation of yohimbine-derivative **52** to yield **54**. Using alternative conditions, *N*-hydroxyindole **55** can be activated by *para*-toluenesulfonyl chloride to enable nucleophilic attack by enamine **56**, affording **57**.[47] Nucleophilic capture can also be achieved through the intermediacy of 3-chloroindolenines, such as the conversion of **58** to **60**.[48] Büchi and co-workers reported an interesting oxidative lactonization of *N*-phthalimido-L-tryptophan (**61**) utilizing DMSO-based oxidants during studies culminating in a total synthesis of tryptoquivaline G (*not shown*).[49] Similar oxidative spirolactonizations have also been effected using excess *N*-bromosuccinimide, with concurrent halogenation of the indole C5 position (*not shown*).[50]

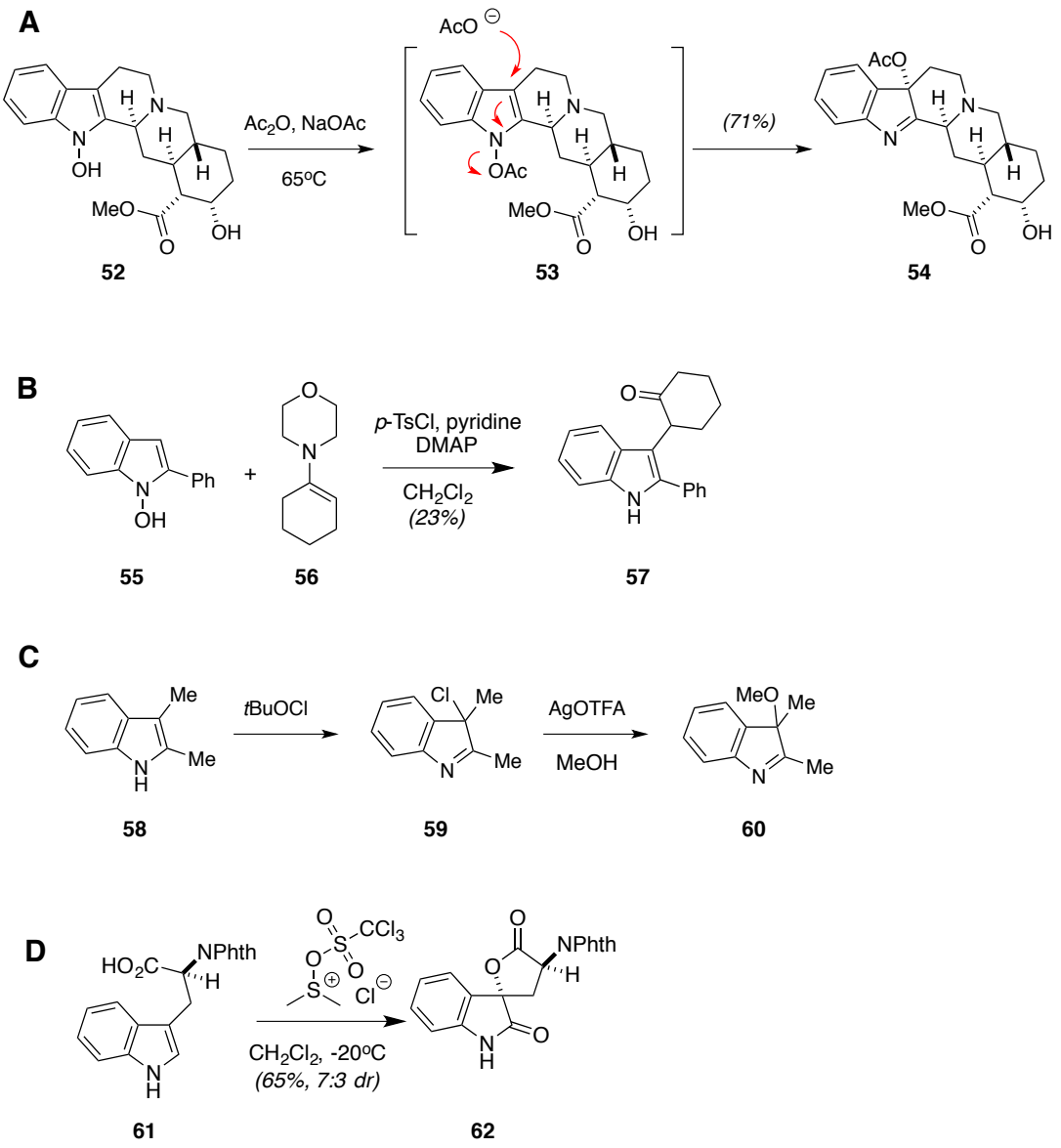


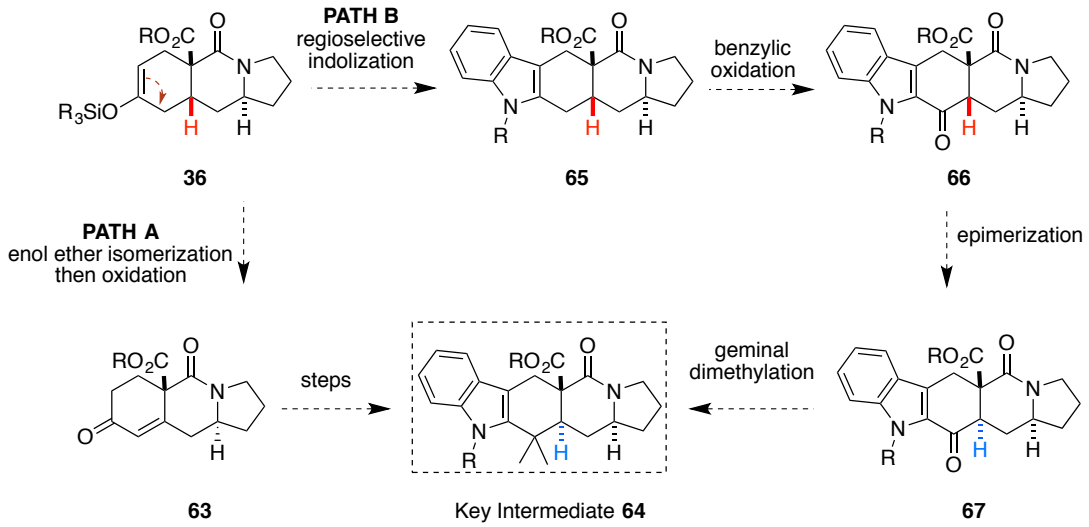
Figure 10: Additional methods to generate electrophilic indole intermediates for C-3 functionalization

Proposed Inversion of Stereochemistry Within Diels-Alder Intermediate 36

As previously described in **Figure 8**, our retrosynthesis entailed rapid construction of the CDE ring system of **3** via a Diels-Alder reaction. This approach would set two of three stereocenters with the correct relative geometry. We projected several ideas to correct the stereochemical configuration with respect to

the C:D ring system of the natural product. Depicted in **Scheme 3** are two hypothesized synthetic routes of how **36**, arising from a Diels-Alder cycloaddition could be converted into core structure **64** with the correct stereochemical configuration. In one possible pathway (Path A), the isomerization of enol ethers within a *cis*-decalin framework could be exploited. Based on previous literature and studies in our own laboratory, the silyl enol ether of **36** may undergo isomerization to produce a thermodynamically-favored enol ether regioisomer.[51,52,53] This would set the stage for regioselective oxidation of the enol ether to produce enone **63**, with selective eradication of the stereocenter at position-4 (highlighted in red). Subsequent stereoselective reduction of the resulting enone and further functionalization may lead to an intermediate such as **64**.

Alternatively, it was hypothesized that the correct stereochemical configuration could be attained through epimerization of the highlighted stereocenter *via* an enol/enolate intermediate (Path B). This could be achieved by initial installation of an indole within the tricyclic system, producing linear pentacycle **65**. Subsequent benzylic oxidation, using one of several known regioselective methods, would afford ketone **66**. This ketone could then be utilized for simple acid or base-mediated epimerization to produce **67**. The ketone would also serve as a functional handle for synthesis of the necessary geminal dimethyl group within **64**.



Scheme 3: Alternative proposed sequences for conversion of **36** to key intermediate **64**

One can qualitatively rationalize the comparative thermodynamic stability of olefins within fused ring systems by understanding their distortional impact on dihedral angles about the C-C bond at the ring junction. Introduction of an olefin within one ring of the decalin system affects the dihedral angles of the fused rings, and overall ring strain is dependent on the placement of this unsaturation. Additionally, the favored placement of unsaturation within one ring of the system varies if the ring system is *cis*-fused or *trans*-fused.[54] To clarify, **Figure 11** seeks to explain the relationship of ring-juncture dihedral angles about the common bond for each type of ring fusion. This figure is interpreted from a review on steroid total synthesis and reactivity by Velluz and co-workers.[55] In the case of a *cis*-decalin framework (**68**), increase of the dihedral angle for the A-ring has an equal distortion effect on the B-ring, thereby bringing the two rings in closer proximity (*i.e.* increasing the concavity). Conversely, in the *trans*-decalin framework (**69**),

increasing the dihedral angle for the A-ring, has an equal and *opposite* effect on the dihedral angle for the B-ring, effectively flattening the B-ring.

In the case of a *cis*-ring junction, placement of unsaturation between the C2-C3 bond of ring-A (**70**) increases the dihedral angles for both A- and B-rings about the central bond, thereby pushing the two rings in closer proximity. This increases the steric strain between the C4 methylene with the axial substituents at C6 and C8. Alternatively, if unsaturation is placed between the C3- C4 bond (**71**), this creates two stabilizing interactions. First, the olefin causes a decrease in the dihedral angles for both ring-A and ring-B, thereby decreasing the concavity of the system (and decreasing steric interactions between axial substituents). Second, elimination of the axial hydrogen at C4 further reduces the steric interaction with C6 and C8.

The explanation for difference in olefin stability in the case of a *trans*-decalin system with unsaturation is more subtle than the *cis*-decalin case. In this instance, the difference in olefin stability between C2-C3 versus C3-C4 is almost negligible if the substituent at C9 is hydrogen. However, if there is a larger substituent at C9, then unsaturation between C3-C4 (**74**) will distort the junction dihedral angles to shorten the distance between the C9 substituent and axial substituents at C5 and C7, thereby increasing $A_{1,3}$ steric interactions. Therefore, unsaturation is thermodynamically favored between C2-C3 instead of C3-C4; contrary to the observed favorability in the *cis*-decalin case.

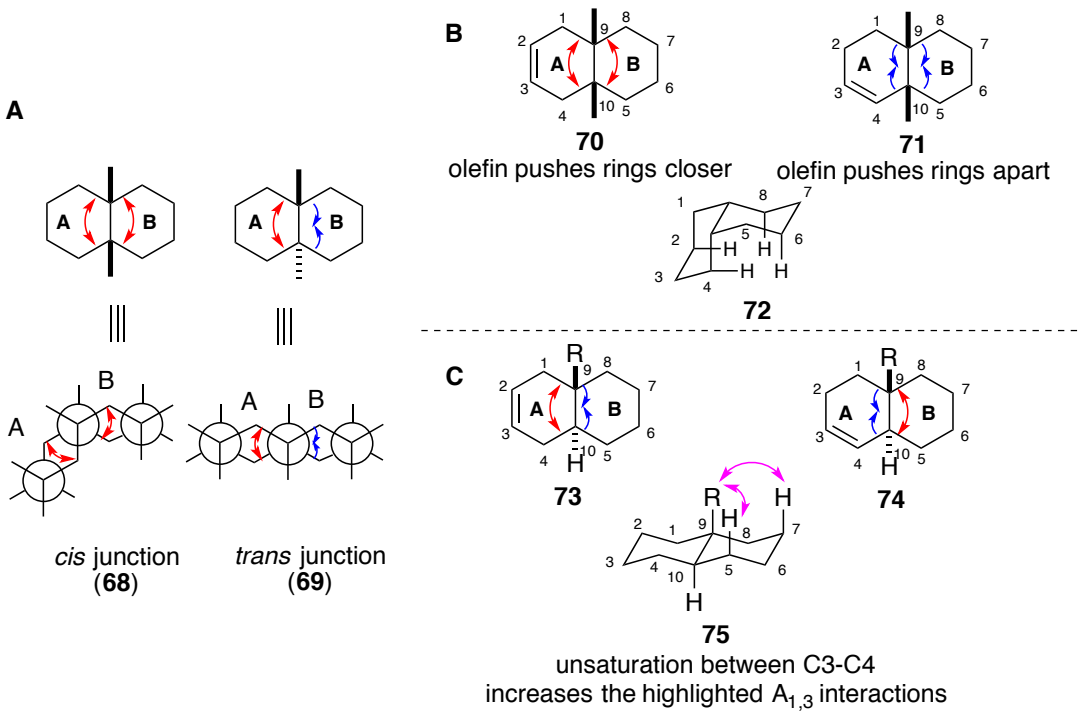


Figure 11: (**A**) Effect of dihedral angle distortion for *cis* and *trans*-fused decalin ring systems (**B**) Effect of unsaturation within *cis* fused ring systems (**C**) Effect of unsaturation within *trans* fused ring systems.

There are many examples, particularly derived from steroid chemistry, which demonstrate the dramatic effect of olefin stability within *cis* or *trans*-fused decalins. The imparted strain affects chemical reactivity and product outcomes, and these trends are related to the challenges associated within this project. **Figure 12** highlights pertinent literature cases. In 2007, Christoffers and co-workers reported the remarkable regioselectivity of the Fischer indole synthesis within fused ring systems of varying stereochemistry.[56] *Cis*-fused **76** yields exclusively "angular" indole product **77**, while *trans*-fused **79** affords exclusively the "linear" indole **80**. Mechanistically, this selectivity may be attributed to the favored enamine intermediate between C3-C4 for **78** (in the case of *cis*) versus the favored enamine intermediate between C2-C3 for **81** (in the case of *trans*). The reactivity of these

simple decalin systems is mirrored by the more complex case of the Fischer indole synthesis between phenylhydrazine with 5β-cholestan-3-one versus 5α-cholestan-3-one, which also display high regioselectivity based on the ring fusion stereochemistry (*not shown*).[57] It should be noted that this reactivity trend, although reliable especially in the case of rigid steroidal structures, is not always general for less rigid decalone structures.[54,58] A notable exception to this selectivity is the Fischer indolization of **82** as described by Stork *en route* toward aspidospermine, with both the *cis* and *trans* ring configurations giving rise to the same observed regiochemistry (**83**).[59]

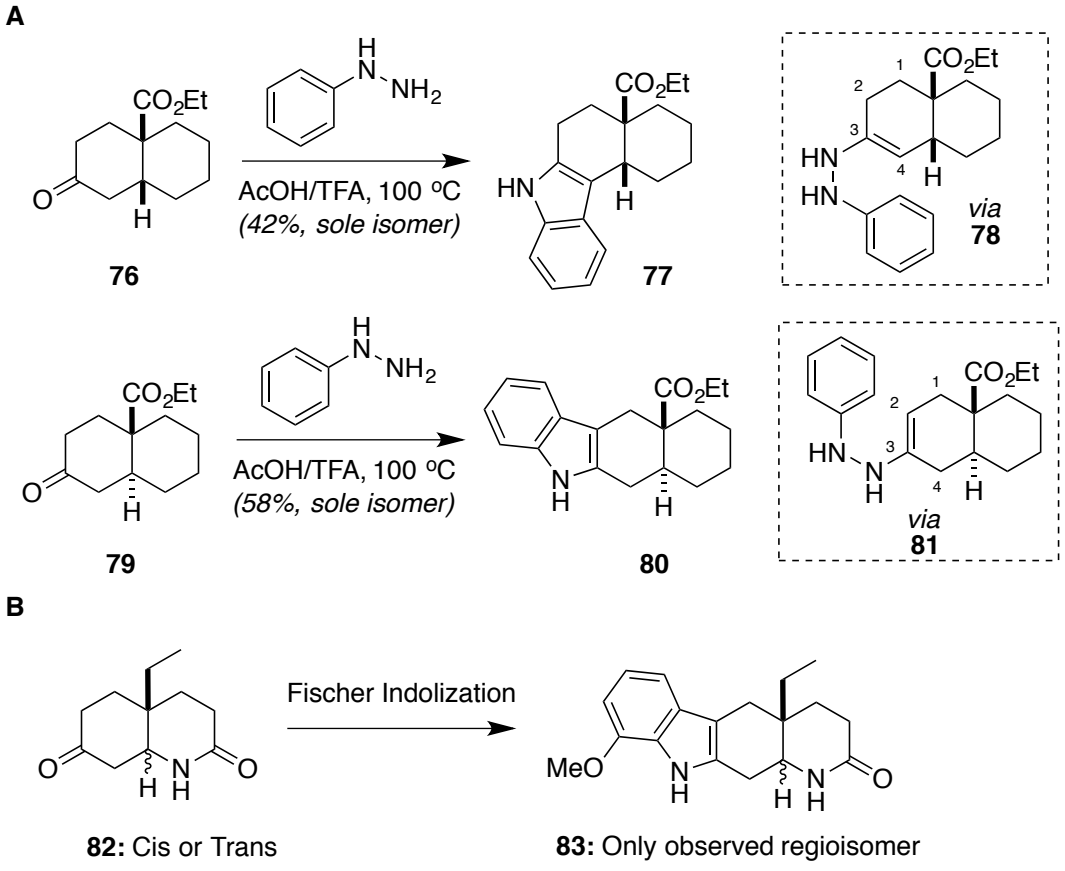


Figure 12: (**A**) Christoffers' study of regiodivergent Fischer indole syntheses in *cis*- versus *trans*-fused ring systems. (**B**) Fischer indolization *en route* toward aspidospermine.

Our laboratory has exploited this difference in olefin stability in the context of total synthesis and in the synthesis of complex enone systems, which would be arguably difficult to attain through other means. During the course of studies on the total synthesis of peribysin E (*not shown*), an interesting olefin isomerization during the key Diels-Alder reaction step was disovered.[51,52] The Diels-Alder reaction between silyloxy diene **84** and (*S*)-Carvone (**85**) followed by treatment with silica gel afforded thermodynamically-favored enol ether **86** in an 11:1 isomerized:non-isomerized product ratio in 70% overall yield. Subsequent Saegusa oxidation with stoichiometric $Pd(OAc)_2$ afforded enone **87**. The realization of this reaction sequence has far-reaching synthetic value, as **87** appears to be the product of a classical Robinson annulation. However, it is noteworthy that the stereochemical relationship between the angular methyl and the distal propenyl group in **87** is the *opposite* configuration to be expected from a Robinson annulation approach. Conversely, the hypothetical Robinson annulation between an enone of type **89** with methyl vinyl ketone (**88**) would be **90**, with an *anti* stereochemical configuration as highlighted, arising from the initially favored axial attack during the Michael Addition step of the annulation.[60,61] Therefore, we have termed enone derivatives of this general sequence to be "iso-Robinson annulation" products; this overall approach is complementary to that of the Robinson annulation.

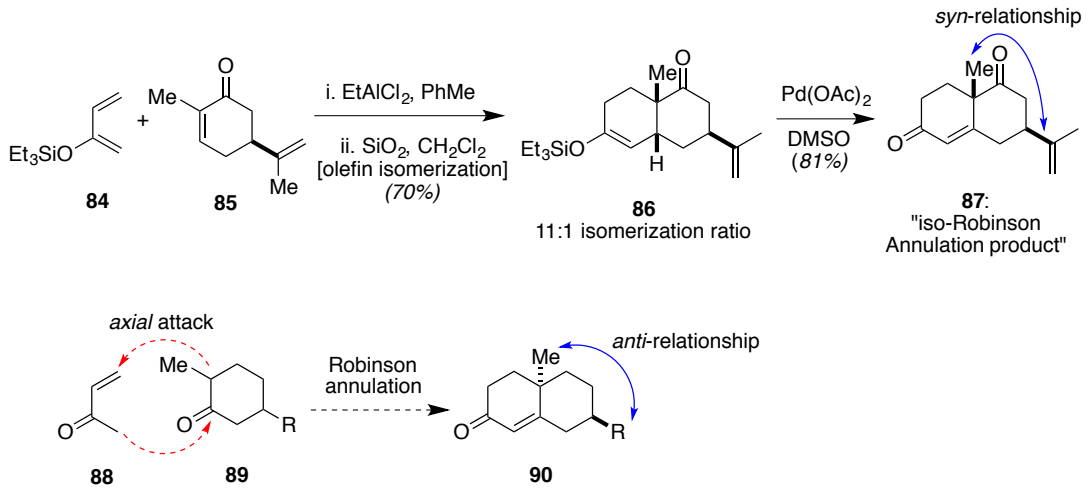


Figure 13: Diels-Alder/Isomerization/Oxidation sequence provides controlled access to stereoisomeric products from that predicted by a classical Robinson annulation approach.

Conclusion:

This chapter seeks to explain our motivation for the study of natural products total synthesis in general, and our specific motivation for the study of aspeverin (**3**). This chapter has also served as an introduction to structurally related natural products and the synthetic challenges these molecules pose, as well as the unique synthetic challenges associated with achieving a total synthesis of aspeverin (**3**). Furthermore, this chapter highlights the literature precedent for the challenging synthetic steps proposed for the synthesis of this target. The following chapter will serve to outline how we ultimately achieved a total synthesis of aspeverin by using our proposed strategies.

References:

1. Butler, M. S. *J. Nat. Prod.* **2004**, *67*, 2141-2153.
2. Williams, D. H.; Stone, M. J.; Hauck, P. R.; Rahman, S. K. *J. Nat. Prod.* **1989**, *52*, 1189-1208.
3. Ji, H.-F.; Li, X.-J.; Zhang, H.-Y. *EMBO Reports* **2009**, *10*, 194-200.
4. Zhang, M.-Q.; Wilkinson, B. *Curr. Opin. Biotech.* **2007**, *18*, 478-488.
5. Schreiber, S. L. *Science* **2000**, *287*, 1964-1969.
6. Xiong, Z.-Q.; Wang, J.-F.; Hao, Y.-Y.; Wang, Y. *Mar. Drugs* **2013**, *11*, 700-717.
7. Charlop-Powers, Z.; Milshteyn, A.; Brady, S. F. *Curr. Opin. Microbiol.* **2014**, *19*, 70-75.
8. Kim, E.; Moore, B. S.; Yoon, Y. J. *Nat. Chem. Biol.* **2015**, *11*, 649-659.
9. Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2016**, *79*, 629-661.
10. Wilson, R. M.; Danishefsky, S. J. *J. Org. Chem.* **2006**, *71*, 8329-8351.
11. Höfle, G.; Bedorf, N.; Steinmetz, H.; Schomburg, D.; Gerth, K.; Reichenbach, H. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1567-1569.
12. Balog, A.; Meng, D.; Kamenecka, T.; Bertinato, P.; Su, D.-S.; Sorensen, E. J.; Danishefsky, S. J. *Angew Chem. Int. Ed. Engl.* **1996**, *35*, 23-24.
13. Su, D.-S.; Meng, D.; Bertinato, P.; Balog, A.; Sorensen, E. J.; Danishefsky, S. J.; Zheng, Y.-H.; Chou, T. C.; He, L.; Horwitz, S. B. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 757-759.
14. Harris, C. R.; Danishefsky, S. J. *J. Org. Chem.* **1999**, *64*, 8434-8456.

15. Rivkin, A.; Yoshimura, F.; Gabarda, A. E.; Cho, Y. S.; Chou, T.-C.; Dong, H.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2004**, *126*, 10913-10922.

16. Rivkin, A.; Chou, T.-C.; Danishefsky, S. J. *Angew Chem. Int. Ed. Engl.* **2005**, *44*, 2838-2850.

17. Nicolaou, K. C.; Vourloumis, D.; Winssinger, N.; Baran, P. S. *Angew. Chem. Int. Ed. Engl.* **2000**, *39*, 44-122

18. Sheehan, J. C.; Henery-Logan, K. R. *J. Am. Chem. Soc.* **1959**, *81*, 3089-3094.

19. Ji, N.-Y.; Liu, X.-H.; Miao, F.-P.; Qiao, M.-F. *Org. Lett.* **2013**, *15*, 2327-2329.

20. Williams, R. M.; Stocking, E. M.; Sanz-Cervera, J. F. *Top. Curr. Chem.* **2000**, *209*, 97-173.

21. Li, S.-M. *Nat. Prod. Rep.* **2010**, *27*, 57-78.

22. Finefield, J. M.; Frisvad, J. C.; Sherman, D. H.; Williams, R. M. *J. Nat. Prod.* **2012**, *75*, 812-833.

23. Xu, W.; Gavia, D. J.; Tang, Y. *Nat. Prod. Rep.* **2014**, *31*, 1474-1487.

24. Sanz-Cervera, J. F.; Glinka, T.; Williams, R. M. *Tetrahedron* **1993**, *49*, 8471-8482.

25. Ding, Y.; Greshock, T. J.; Miller, K. A.; Sherman, D. H.; Williams, R. M. *Org. Lett.* **2008**, *10*, 4863-4866.

26. Miller, K. A.; Williams, R. M. *Chem. Soc. Rev.* **2009**, *38*, 3160-3174.

27. Schkeryantz, J. M.; Woo, J. C. G.; Siliphaivanh, P.; Depew, K. M.; Danishefsky, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 11964-11975.

28. Edmondson, S. D.; Danishefsky, S. J. *Angew. Chem. Int. Ed. Engl.* **1998**, *37*, 1138-1140.

29. von Nussbaum, F.; Danishefsky, S. J. *Angew. Chem. Int. Ed. Engl.* **2000**, *12*, 2175-2178.

30. Li, C.-W.; Wu, C.-J.; Cui, C.-B.; Xu, L.-L.; Cao, F.; Zhu, H.-J. *RSC Adv.* **2016**, *6*, 73383-73387.

31. Mugishima, T.; Tsuda, M.; Kasai, Y.; Ishiyama, H.; Fukushi, E.; Kawabata, J.; Watanabe, M.; Akao, K.; Kobayashi, J. *J. Org. Chem.* **2005**, *70*, 9430-9435.

32. Mercado-Marin, E.V.; Garcia-Reynaga, P.; Romminger, S.; Pimenta, E.F.; Romney, D.K.; Lodewyk, M.W.; Williams, D.E.; Andersen, R.J.; Miller, S.J.; Tantillo, D.J.; Berlinck, R.G.S.; Sarpong, R. *Nature* **2014**, *509*, 318-324.

33. Bian, Z.; Marvin, C. C.; Martin, S. F. *J. Am. Chem. Soc.* **2013**, *135*, 10886-10889.

34. Bian, Z.; Marvin, C. C.; Pettersson, M.; Martin, S. F. **2014**, *136*, 14184-14192.

35. Kong, K.; Enquist Jr., J. A.; McCallum, M. E.; Smith, G. M.; Matsumaru, T.; Menhaji-Klotz, E.; Wood, J. L. *J. Am. Chem. Soc.* **2013**, *135*, 10890-10893.

36. Mundal, D. A.; Sarpong, R. *Org. Lett.* **2013**, *15*, 4952-4955.

37. Guerrero, C. A.; Sorensen, E. J. *Org. Lett.* **2011**, *13*, 5164-5167.

38. Mercado-Marin, E. V.; Sarpong, R. *Chem. Sci.* **2015**, *6*, 5048-5052.

39. Kushida, N.; Watanabe, N.; Okuda, T.; Yokoyama, F.; Gyobu, Y.; Yaguchi, T. *J. Antibiot.* **2007**, *60*, 667-673.

40. Fürst, A.; Plattner, P. A. *Helv. Chim. Acta* **1949**, *32*, 275-283.

41. Awang, D. V. C.; Vincent, A. *Can. J. Chem.* **1980**, *58*, 1589-1591.

42. Braun, N. A.; Bray, J. D.; Ciufolini, M. A. *Tetrahedron Lett.* **1999**, *40,* 4985-4988.

43. Ishikawa, H.; Takayama, H.; Aimi, N. *Tetrahedron Lett.* **2002**, *43*, 5637-5639.

44. Feldman, K. S.; Vidulova, D. B.; Karatjas, A. G. *J. Org. Chem.* **2005**, *70*, 6429-6440.

45. Zhdankin, V. V.; Crittell, C. M.; Stang, P. J. *Tetrahedron Lett.* **1990**, *31*, 4821-4824.

46. Somei, M.; Noguchi, K.; Yamada, F. *Heterocycles* **2001**, *55*, 1237-1240.

47. Nagayoshi, T.; Saeki, S.; Hamana, M. *Chem. Pharm. Bull.* **1981**, *29*, 1920-1926.

48. Gassman, P. G.; Campbell, A.; Mehta, G. *Tetrahedron* **1972**, *28*, 2749-2758.

49. Buchi, G.; DeShong, P. R.; Katsumura, S.; Sugimura, Y. *J. Am. Chem. Soc.* **1979**, *101*, 5084-5086.

50. Patchornik, A.; Lawson, W. B.; Witkop, B. *J. Am. Chem. Soc.*, **1958**, 4747-4748.

51. Angeles, A. R.; Waters, S. P.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2008**, *130*, 13765-13770.

52. Angeles, A. R.; Dorn, D. C.; Kou, C.A.; Moore, M. A.; Danishefsky, S. J. *Angew. Chem. Int. Ed. Eng.* **2007**, *46*, 1451-1454.

53. Peng, F.; Dai, M.; Angeles, A. R.; Danishefsky, S. J. *Chem. Sci.* **2012**, *3*, 3076-3080.

54. Huffman, J. W.; Balke, W. H. *J. Org. Chem.* **1988**, *53*, 3828-3831.

55. Velluz, L; Valls, J.; Nominé, G. *Angew. Chem. Int. Ed. Eng.* **1965,** *4*, 181-200.

56. Diedrich, C. L.; Frey, W.; Christoffers, J. *Eur. J. Org. Chem.* **2007**, *28*, 4731-4737.

57. Ban, Y.; Sato, Y. *Chem. Pharm. Bull.* **1965**, *13*, 1073-1077.

58. Yanagita, M.; Yamakawa, K. *J. Org. Chem.* **1957**, *22*, 291-297.

59. Stork, G.; Dolfini, J. E. *J. Am. Chem. Soc.* **1963**, *85*, 2872-2873.

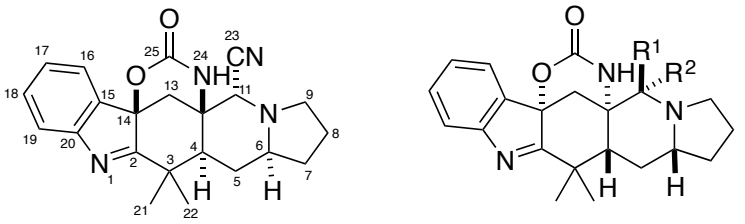
60. Howe, R.; McQuillin, F. J. *J. Chem. Soc.* **1955**, 2423-2428.

61. Wenkert, E.; Berges, D. A. *J. Am. Chem. Soc.* **1967**, *89*, 2507-2509.

Chapter 2: Total Synthesis of Aspeverin and Penicimutamide A

The previous chapter serves as an introduction to explain our motive and strategies for accessing aspeverin (**1**, **Figure 1**) through chemical synthesis. Herein, we describe how retrosynthetic disconnections were realized to achieve a total synthesis of aspeverin **1**. Key to our synthesis is a diastereoselective Diels−Alder cycloaddition to set two of three stereocenters with proper relative configuration. The necessary dienophile could be accessed in two steps from known indolizidine **10** **(Scheme 1)**.[1] Starting with *N*-(*tert*-butyloxycarbonyl)-2-pyrrolidinone **5**, reduction with diisobutylaluminum hydride yielded a hemiaminal, which underwent allylation with allyltrimethylsilane to produce a mixture of **6** and **7** in 88% combined overall yield. Both of these products react productively in the subsequent steps. Acid deprotection of the Boc group, followed by acylation with acryloyl chloride generates **8** in 74% overall yield. The resultant diene efficiently underwent olefin metathesis with Grubbs' second generation catalyst (**9**) at room temperature to afford known indolizidine **10**. Iodination of this compound under modified conditions reported by Johnson and co-workers afforded **11**.[2] This iodide was subjected to Pd-catalyzed carbonylation to yield **12**. As anticipated, **12** underwent a Diels−Alder reaction with silyloxydiene **13** using $ZnCl_2$ as a Lewis acid, affording **14** in 91% yield. Facial selectivity here is guided by the distal stereocenter within **12**, giving rise to a single diastereomer as anticipated, with correct stereochemistry at positions 6 and 12 (according to the natural product numbering described in

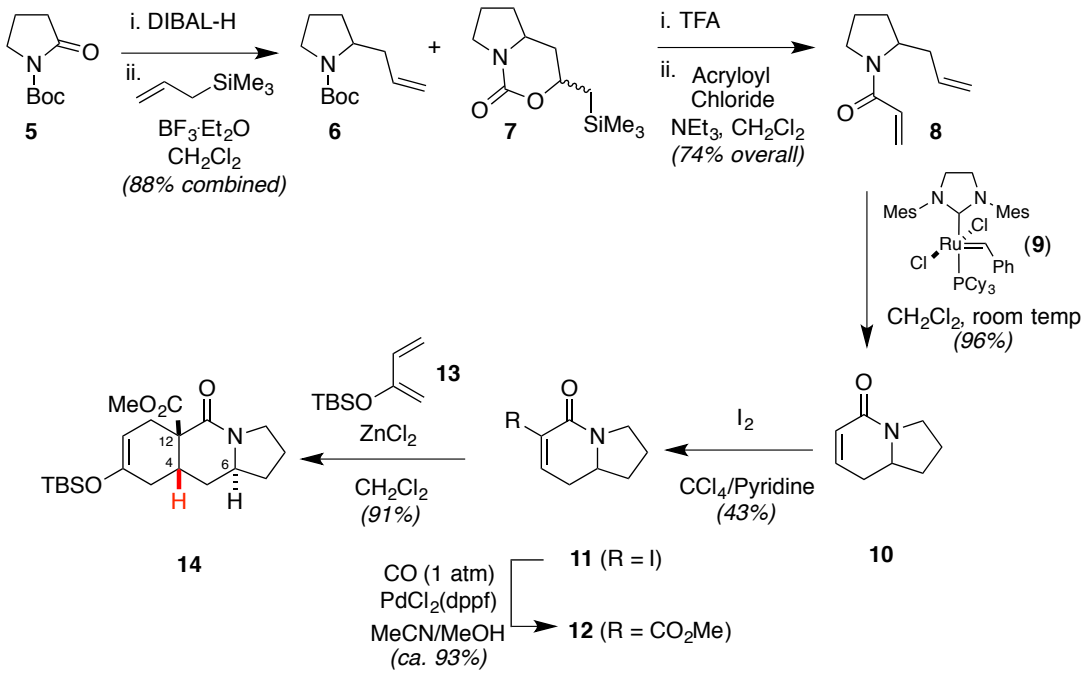
Figure 1).[3] Notably, the stereochemistry at position 4 is incorrect, but was resolved in a later sequence.



aspeverin (**1**)

penicimutamides A-C
A: (**2**, $R^1 = R^2 = O$)
B: (**3**, $R^1 = OH$, $R^2 = H$)
C: (**4**, $R^1 = H$, $R^2 = H$)

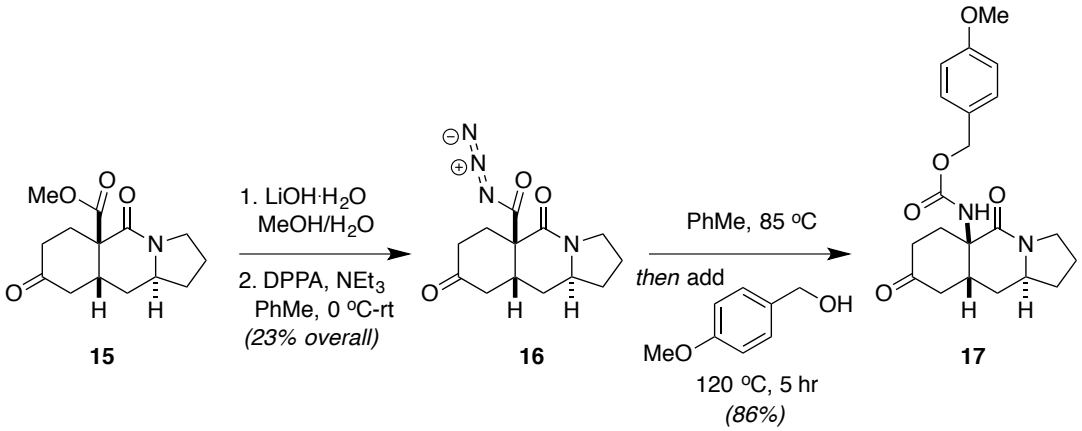
Figure 1: Structures of carbamate-containing indole alkaloids aspeverin (**1**) and penicimutamides A-C (**2-4**)



Scheme 1: Synthesis of the C:D:E ring system of aspeverin via a diastereoselective Diels−Alder reaction between **12** and **13**.

Prior to this resolution, we wanted to test the viability of a Curtius rearrangement at the C:D ring juncture. We had concern regarding the angular β-

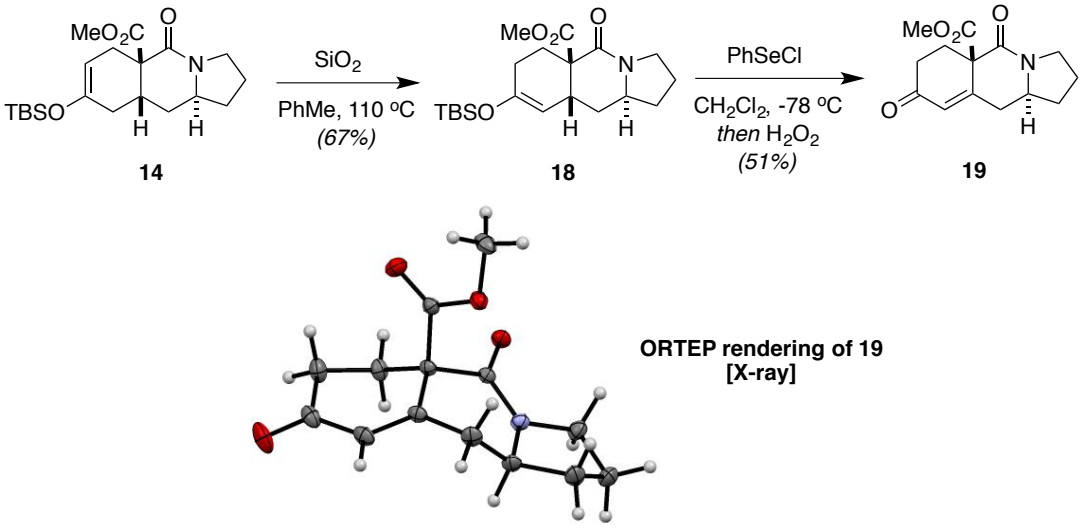
amido ester, which could foreseeably undergo undesired decarboxylation upon conversion to the carboxylic acid. **Scheme 2** shows our model to install the angular carbamate. Fortunately, even without optimization, no apparent decarboxylation was observed upon saponification of **15** at room temperature. This crude acid was converted to acyl azide **16** in 54% yield upon treatment with diphenylphosphoryl azide (DPPA).[4] Remarkably, this intermediate was stable likely due to steric hindrance, and could be conveniently isolated and characterized following purification by silica gel chromatography. Thermolysis at 85 ^{o}C fully converted **16** to the corresponding isocyanate, at which point *para*-methoxybenzyl alcohol was added to generate carbamate **17** in 86% yield. Installation of the carbamate was ultimately planned to take place later in the synthesis, although these preliminary experiments established the feasibility of the desired transformation.



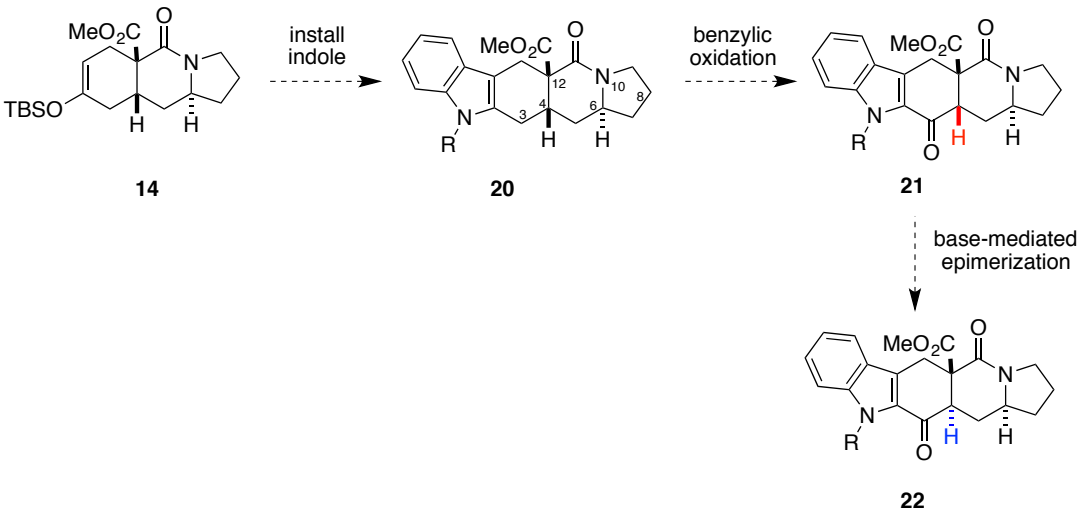
Scheme 2: Successful Curtius Rearrangement sequence starting from **15**.

Moving forward, we sought to access the correct stereochemical configuration at position 4. In one strategy, we were able to isomerize enol ether **14** using a method previously reported by our lab (**Scheme 3**).[5] Prolonged heating of the compound in the presence of silica gel, with reaction monitoring by ^{1}H NMR,

furnished desired olefin regioisomer **18** in 67% isolated yield with a favorable 10:1 selectivity (shown isomer). Subsequent phenylselenation followed by elimination of the corresponding *in situ* generated selenoxide afforded enone **19** in 51% overall yield. Notably, this compound contains the opposite stereochemical configuration one would expect to arise from a Robinson annulation strategy. Preliminary attempts to reduce this "iso-Robinson annulation" product via hydrogenation using various solvents generally resulted in low diastereoselectivity. Moreover, initial attempts at dissolving metal reduction or γ-enolate alkylation of this enone provided complex mixtures. Although the conditions attempted to functionalize or reduce this enone were not exhaustive, we explored an alternatively described synthetic route (**Scheme 4**) and diverted attention upon promising initial success. We therefore focused our efforts toward the natural product through this alternative sequence. In this route, we hoped to install the indole ring prior to epimerization of the C:D ring junction. Additionally, we would use the inherent reactivity of the indole heterocycle to functionalize benzylic position 3.



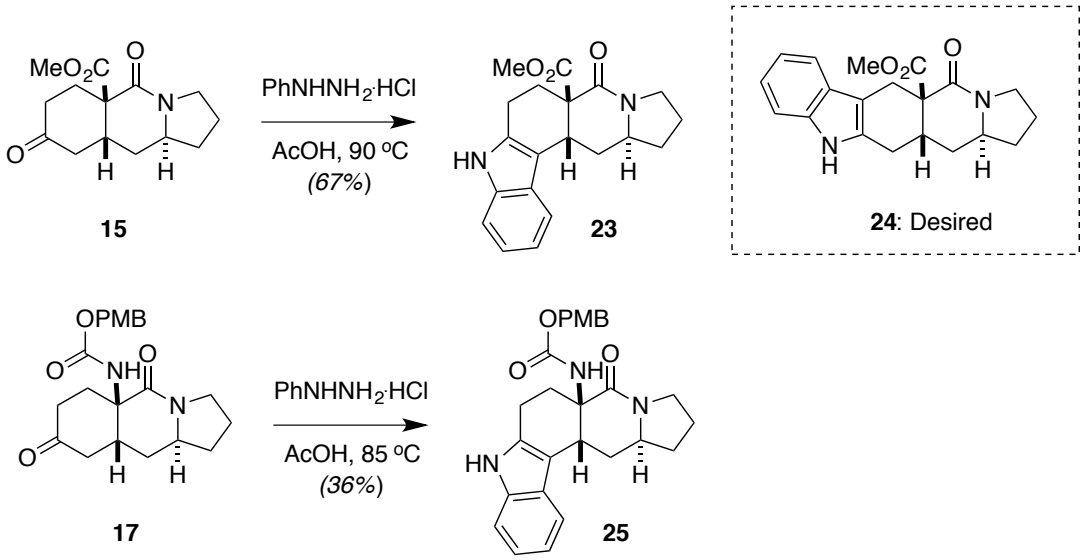
Scheme 3: Synthesis of enone **19**, and X-ray structure confirming relative stereochemistry



Scheme 4: An alternative synthetic route to **22**.

In this regard, initial attempts at indole installation relied on Fischer indole syntheses from ketones **15** or **17** (**Scheme 5**).[6] Treatment of **15** with phenylhydrazine in AcOH afforded undesired indole derivative **23** in 67% isolated yield, with only trace formation of the desired regioisomer **24**. Similarly,

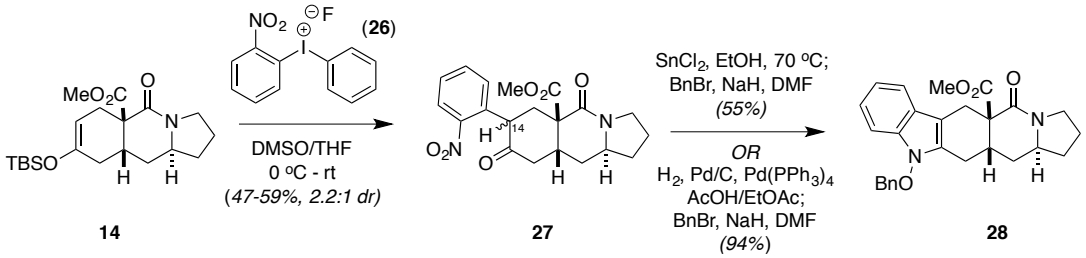
indolization of ketone **17** again afforded a single indole product, with isolation of only the undesired regioisomer **25**, albeit in low yield. These results are unsurprising in retrospect, given the favored regiochemistry of Fischer indolization within *cis*-fused decalin systems as discussed in the previous chapter.



Scheme 5: Fischer indolization of **15** and **17** with incorrect regioselectivity.

Fortunately, we found that regioselective installation of the desired indole ring system could be accomplished using *o*-(nitrophenyl)phenyliodonium fluoride (NPIF, **26**), a reagent developed by Rawal and co-workers to address similar issues encountered with regard to Fischer indole regioselectivity (**Scheme 6**).[7] Treatment of **14** with this electrophilic reagent affords **27** as a single regioisomer at position 14 and as a 2.2:1 mixture of inconsequential diastereomers in 47-59% isolated yield. This intermediary nitroaryl ketone conveniently undergoes reductive cyclization. Several initial conditions attempted for reduction including catalytic hydrogenation and Fe^0/AcOH afforded mixtures of the desired indole as well as the partially-

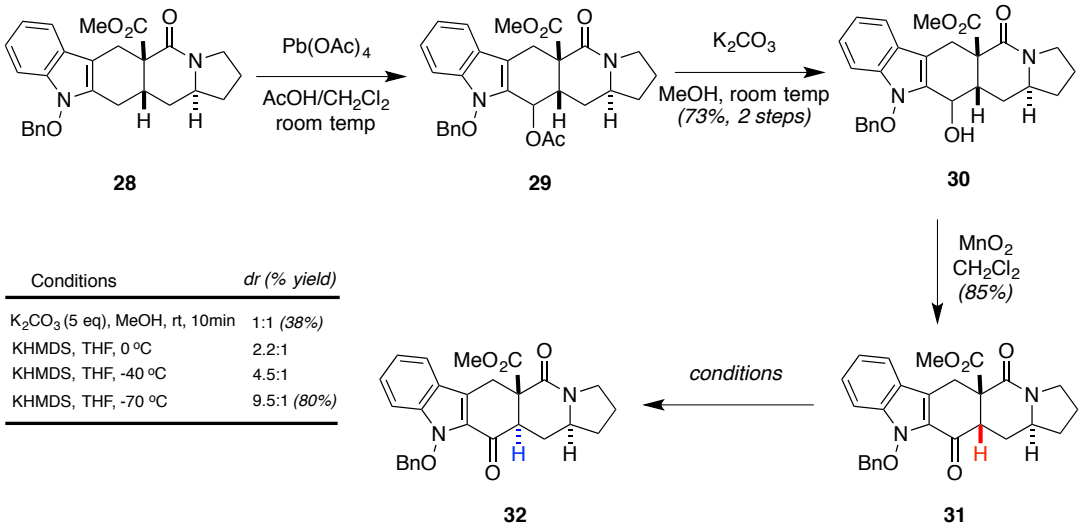
reduced *N*-hydroxyindole derivative. Ultimately, it was found that **27** could be exclusively converted to the *N*-hydroxyindole either upon treatment with $SnCl_2$ or upon reduction by H_2 and Pd/C in the presence of $Pd(PPh_3)_4$ as described by Belley and co-workers.[8,9] The *N*-hydroxyindole itself was not isolated; rather it was directly alkylated upon treatment with NaH and benzyl bromide. The latter described hydrogenation method was found to be higher yielding, without generation of tin byproducts, and afforded *N*-benzyloxyindole **28** in 94% overall yield. According to Belley and co-workers, $Pd(PPh_3)_4$ serves to poison the hydrogenation catalyst, such that cyclization of the nitroso intermediate is faster than further reduction. Notably, this *N*-benzyloxyindole was carried forward with the thought that the higher oxidation state could actually simplify later steps, wherein the angular carbamate would directly cyclize onto the indole core in a fashion similar to that described by Somei and co-workers for *N*-hydroxyindoles.[10]



Scheme 6: Regioselective electrophilic arylation using NPIF (**26**) and reductive cyclization to *N*-benzyloxyindole **28**.

With the desired fused indole ring system in place, we explored means to epimerize position 4 via the installation of a benzylic ketone. Several oxidative conditions are known to regioselectively functionalize at the desired benzylic position of the indole.[11-13] After some experimentation, it was found that *N*-

benzyloxyindole **28** underwent highly efficient α-acetoxylation upon treatment with lead (IV) acetate in acetic acid according to the procedure of Kametani and co-worker (**Scheme 7**).[14] The resultant crude acetate was subjected to transesterification with K_2CO_3 in MeOH to efficiently yield alcohol **29** in 73% overall yield as a single, unassigned diastereomer. Further oxidation upon treatment with excess MnO_2 afforded the desired ketone in 85% yield. As anticipated, epimerization of the α-stereocenter was accomplished under basic conditions. Brief exposure to K_2CO_3 in methanol yielded a 1:1 diastereomeric ratio of separable epimers, with isolation of **32** in 38% yield. Fortunately, enolization of **31** with KHMDS at low temperature and quenching at −70 °C with aqueous NH_4Cl afforded **32** in a satisfactory 9.5:1 *dr* and 80% isolated yield.

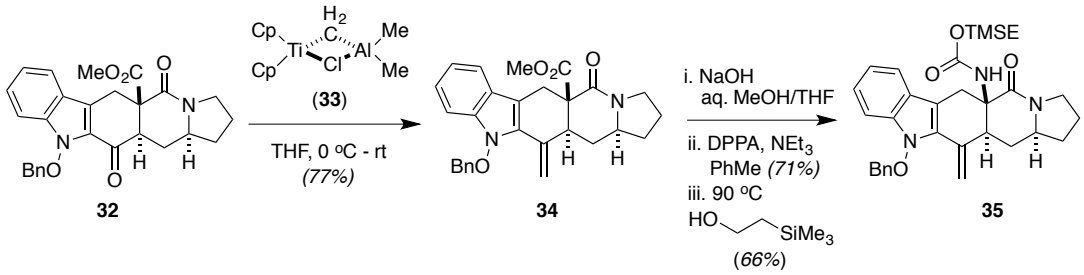


Scheme 7: Benzylic oxidation and successful epimerization to yield **32**

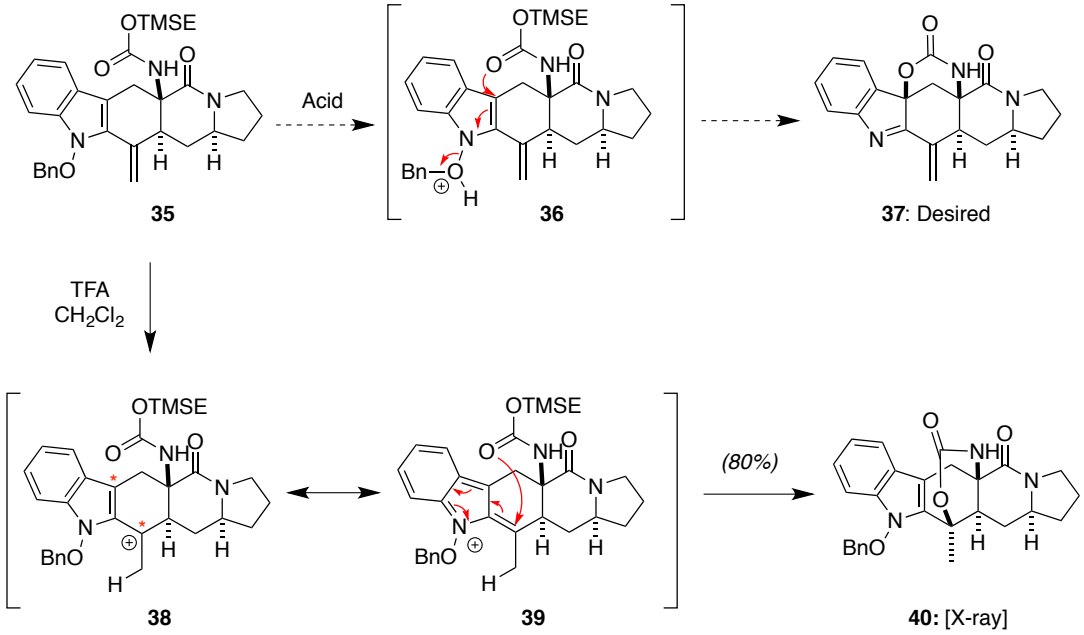
With the successful epimerization at carbon 4 within pentacyclic ring system **32**, we were poised to complete functionalization of the natural product core

scaffold. We turned our attention toward methods to convert: 1) the ketone functionality into the requisite geminal dimethyl group, and 2) the angular ester into the requisite carbamate via Curtius rearrangement. Due to the potential of epimerization at the α-stereocenter under basic conditions, we opted to functionalize the ketone prior to ester saponification. We found that this ketone could be conveniently functionalized as an olefin upon treatment with the Tebbe reagent (**33**) to afford methylidene **34** (**Scheme 8**).[15] Subsequent ester hydrolysis afforded the desired acid, which was subjected Curtius rearrangement conditions. Formation of the acyl azide with diphenylphosphoryl azide, followed by thermolysis in the presence of 2-trimethylsilylethanol afforded the desired carbamate **35**. At this point, we were posed with two main challenges to tackle: 1) cyclization of the formed carbamate onto the C-3 carbon of the indole ring, and 2) conversion of the exocyclic olefin to a geminal dimethyl group. Notably, despite literature precedent for strategies to accomplish the latter goal, this still remained a non-obvious task.[16,17] Perhaps somewhat naïvely, it was envisioned that acid treatment of **35** could render the indole core susceptible toward nucleophilic attack. In this way, the carbamate may cyclize onto the indole ring with loss of benzyl alcohol to produce **37**. Alternatively, protonation of the exocyclic olefin would generate a resonance-stabilized tertiary carbocation **38**, with two potential electrophilic sites (highlighted with asterisks within intermediate **38**). Although the anticipated transformation did not occur, it was found that upon exposure of this compound to a slight excess of trifluoroacetic acid, the carbamate readily adds across the olefin, cleanly affording cyclic carbamate **40** in 80% isolated yield.[18] The structure of this molecule was

further confirmed with a preliminary single molecule X-ray crystal structure (**Figure 2**). It is noteworthy that this molecule is isomeric to the bridging carbamate found within the natural product. This further corroborates the projected nucleophilic capacity of the carbamate functionality and the possibility to generate such constrained, yet stable molecular architectures.



Scheme 8: Ketone olefination and Curtius rearrangement



Scheme 9: Acid-mediated carbamate cyclization to generate **40**

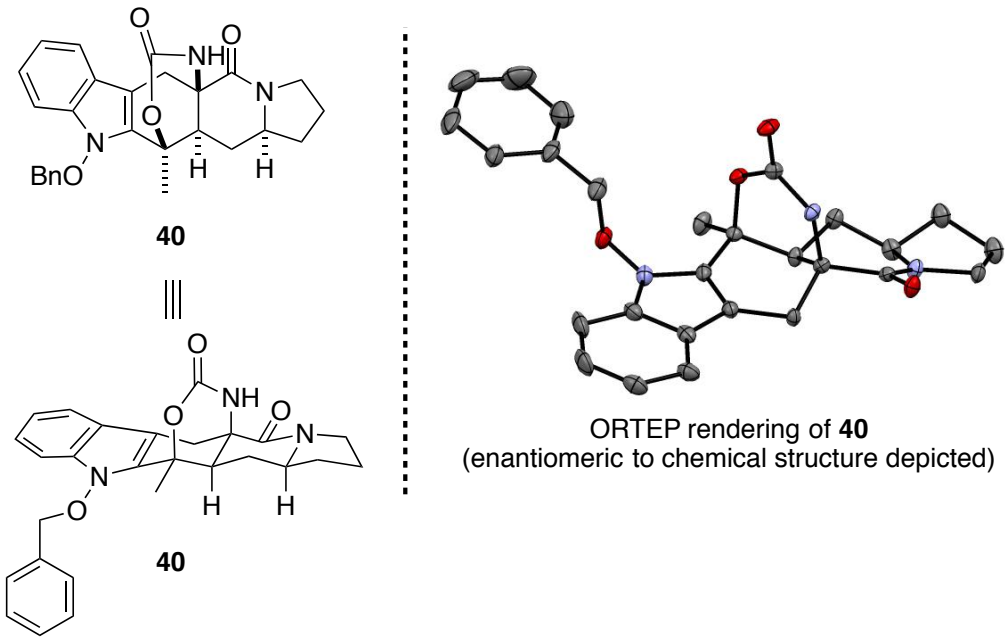
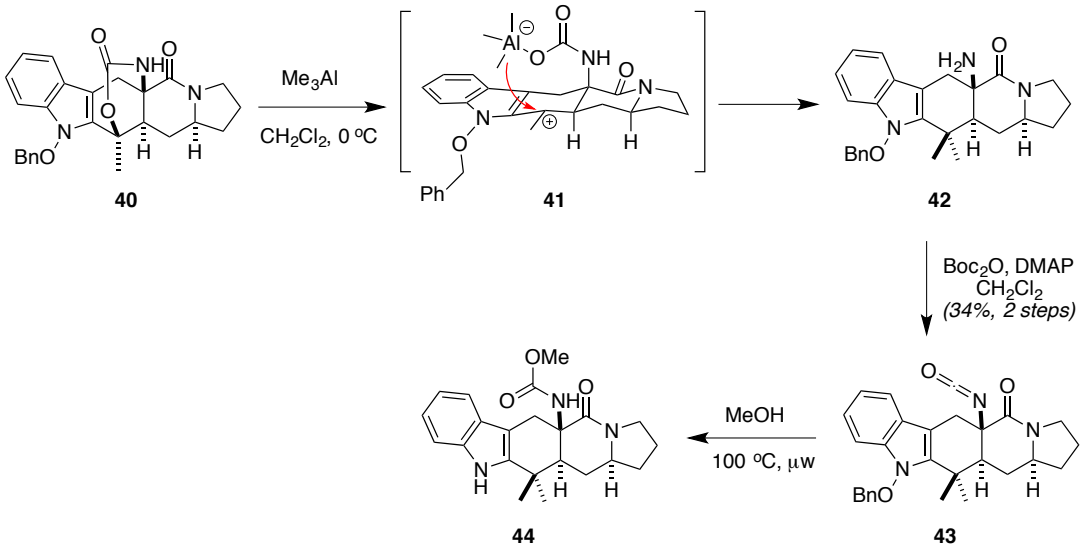


Figure 2: ORTEP rendering of the X-ray crystal structure for **40**

Fortuitously, identification of structure **40** inspired a solution to the synthetic challenge of installing a geminal dimethyl group at position 3. With a single methyl group installed and a leaving group in place (*i.e.* benzylic carbamate), it was hypothesized that activation of cyclic carbamate **40** with a Lewis acid may serve to once again generate an intermediary stabilized tertiary carbocation. Moreover, if the carbocation was formed in the presence of a methyl nucleophile source, the desired geminal dimethyl group could be installed. This strategy was deemed successful, with trimethylaluminum filling both of these criteria. Indeed, upon exposure of **40** to excess Me_3Al, the desired geminal dimethyl group was formed and primary amine **42** was isolated (**Scheme 10**). This reaction is presumed to occur via the resonance-stabilized carbocation intermediate **40**. Although this compound could be purified for characterization, it was carried forward crude in a second step involving

treatment with di-*tert*-butyl dicarbonate (Boc_2O), which generated stable isocyanate **43** rather than the expected *tert*-butyloxycarbamate in 34% yield over two steps. Di-*tert*-butyl dicarbonate has been previously reported as a reagent for the synthesis of isocyanates, especially from aromatic and hindered amine substrates.[19] Nonetheless, heating this isocyanate **43** in methanol generated methyl carbamate **44**. Moreover, although methanolysis is relatively fast (within approximately 1 hour), it was noted that an inseparable side-product formed under these reaction conditions corresponding to the reduced indolic form. Prolonged heating (*ca.* 13 hours) fully converted **43** to free indole **44** with liberation of benzaldehyde, based on ^{1}H NMR analysis of the crude reaction mixture. This process is mechanistically unclear, but literature precedent suggests that similar *N*-benzyloxy derivatives undergo disproportionation either via heterolytic N-O bond cleavage or retro-ene type mechanisms (**Figure 3**).[20,21,22] Similarly, Wu and co-workers demonstrated that pyridine *N*-oxide (**53**) can displace benzylic halides to produce intermediates of type **54**, which undergo disproportionation via a mechanism similar to the Kornblum oxidation.[23,24]



Scheme 10: Installation of a geminal dimethyl group and synthesis of **44**.

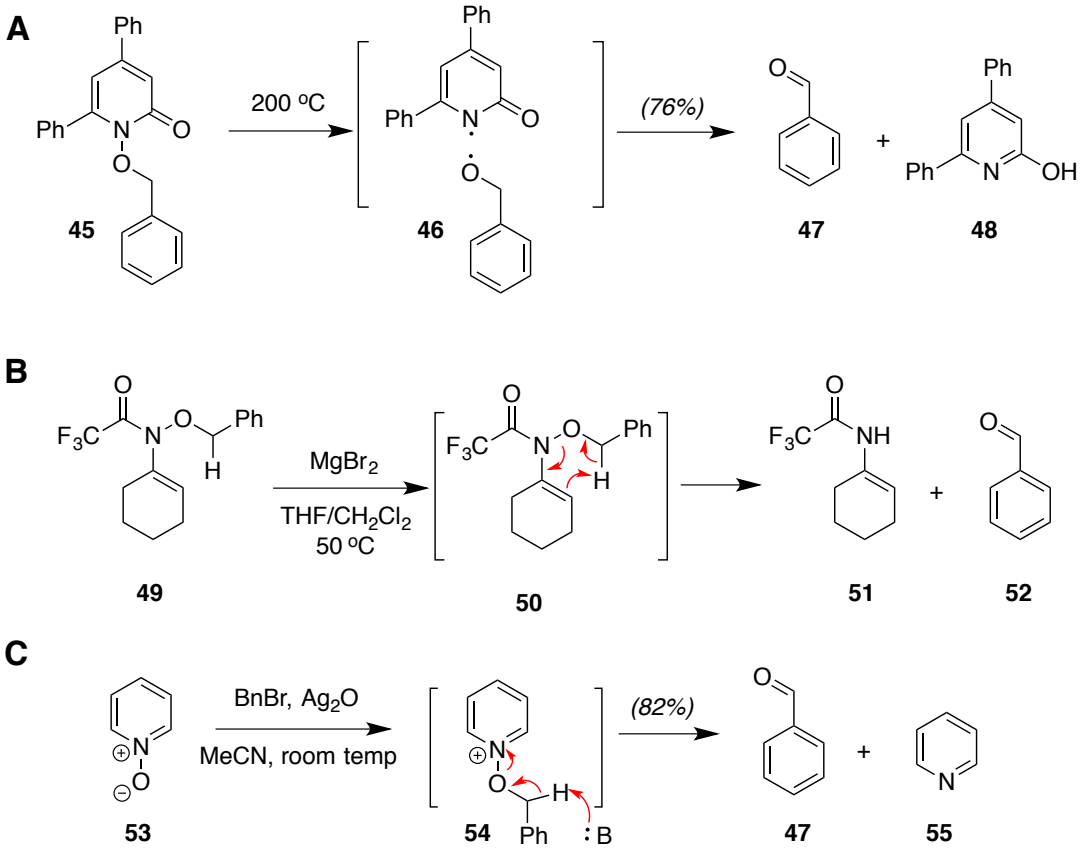
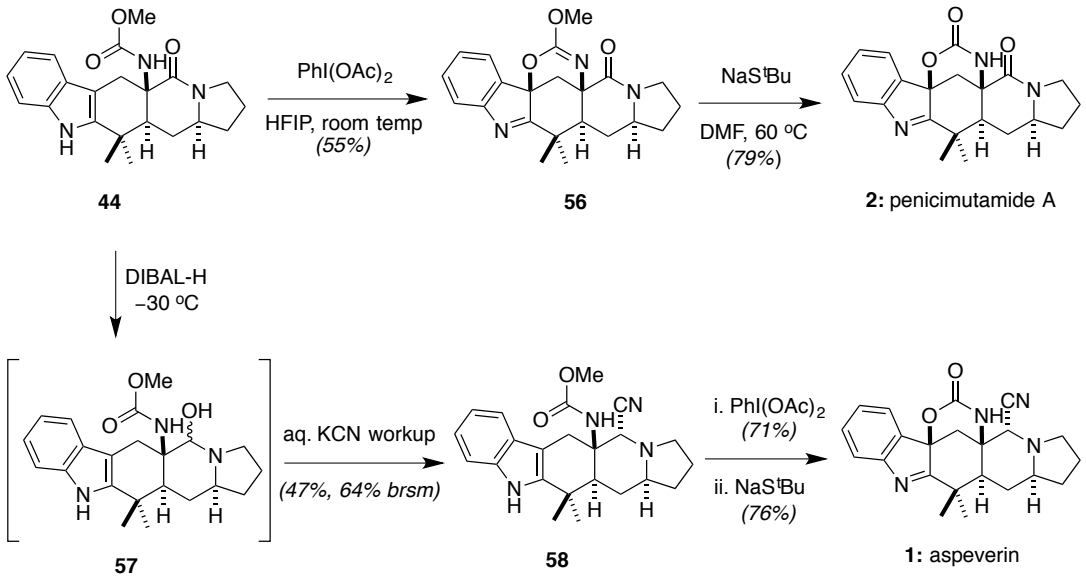


Figure 3: Literature precedent for observed disproportionation of **43** upon thermal methanolysis (**A**) disproportionation of N-benzyloxypyridone **45** (**B**) retro-ene reaction of *N*-benzyloxyenamine **49** (**C**) pyridine *N*-oxide oxidation of benzylic halides

With a functionalized scaffold of aspeverin (except for the α-cyanoamine) in hand, attempts at oxidative carbamate cyclization were undertaken. Surprisingly, this oxidative cyclization was realized rather quickly upon screening of a few reaction conditions. It was envisaged that development of this transformation would be rather arduous. Nonetheless, a previously reported protocol effected the desired reactivity.[25,26] Indeed, treatment of **44** with $PhI(OAc)_2$ in 1,1,1,3,3,3-hexafluoroisopropanol (HFIP) produced indolenine **56** in 55% isolated yield (**Scheme 11**). Interestingly, the methylimidate proved stable to these conditions, but could be subsequently dealkylated with sodium 2-methyl-2-propanethiolate in DMF to afford **2**. Moreover, during the course of writing this thesis, Li and co-workers isolated **2** from a mutant strain of *Penicillium purpurogenum* G59, and named this molecule penicimutamide A.[27] Our ^{1}H and ^{13}C spectral data in retrospect were in full agreement with that reported. The penultimate functionalization required was to install the α-cyanoamine, which was envisioned to arise from reduction of **44** to afford an intermediate hemiaminal (**57**) that should undergo a thermodynamic exchange upon workup with cyanide. Treatment with DIBAL-H followed by workup with aqueous potassium cyanide afforded α-cyanoamine **58** in 47% yield as a single stereoisomer (64% based on recovered starting material).[28] Subsequent treatment with $PhI(OAc)_2$ as described before afforded the cyclic methyl carboximidate in 71% yield, which similarly underwent nucleophilic demethylation with sodium 2-methyl-2-propanethiolate to afford aspeverin **1** in 76% yield. Spectroscopic data was in agreement with that reported by Ji and co-workers for the isolated natural product.[29]



Scheme 11: Oxidative carbamate cyclization and completion of aspeverin (**1**) and penicimutamide A (**2**)

Conclusions and Future Projections:

The chemistry described in this chapter delineates the first and only total syntheses of aspeverin (**1**) and penicimutamide A (**2**) to date.[30] Our synthetic route toward this molecule features several unique synthetic transformations, which should have further utility in the synthesis of other complex natural products. We were able to construct the core scaffold via a highly regio- and diastereoselective Diels−Alder reaction. We further demonstrated that an incorrectly set stereocenter could be inverted with a high level of diastereoselectivity via base-mediated epimerization of an intermediary ketone. Notably, this ketone later served as a functional handle for geminal dimethylation via carefully devised synthetic manipulations that hinge upon the exploitation of a presumed tertiary carbocation. An olefin hydration − ring opening − alkylation sequence ultimately proved successful in providing the desired geminal dimethyl group. Finally, we

demonstrated that the bridging carbamate unique to this natural product could be forged via an iodine(III)-mediated oxidative cyclization of the pendant carbamate as was initially planned in our retrosynthetic proposal. With several established strategies to access the functionalized core of aspeverin, it could be beneficial to explore a diversification platform to provide access to structurally related molecules.

Future Projections:

We hope our synthetic route can provide insight toward the synthesis of related indolic natural products. One set of fungal indolic alkaloids in particular, PF1270A-C (**59-61**) isolated from *Penicillium waksmanii*, shares many of the unique architectural features of aspeverin, but with added synthetic challenges (**Figure 4**).[31] PF1270A (**59**) is redrawn as its enantiomer for clarity and to exemplify its structural similarities with aspeverin (**1**). Notably, these three molecules possess high affinity for human histamine H3 receptors (H3Rs) (K_i = 0.047, 0.12, and 0.22 μM respectively) and act as potent agonists for these receptors. H3Rs have been implicated as targets in a variety of disease types, which include obesity and diabetes, and several central nervous system (CNS) diseases including sleeping disorders, schizophrenia, movement disorders, and attention deficit hyperactivity disorder (ADHD).[32]

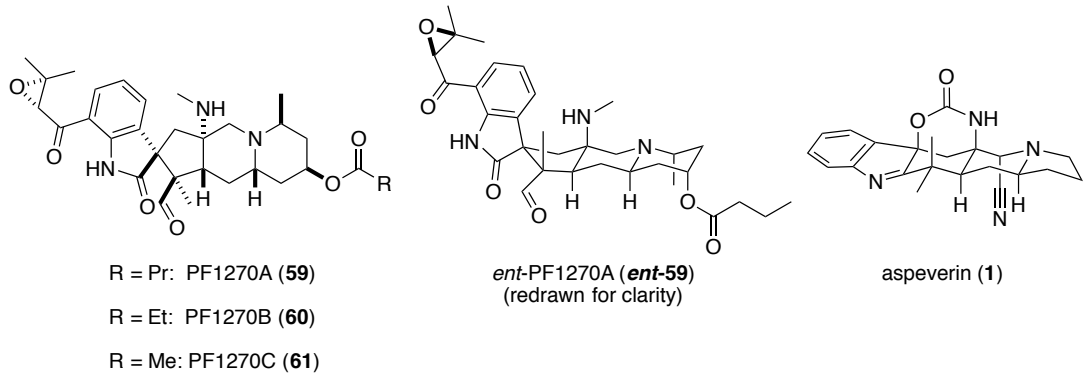


Figure 4: Structures of PF1270A-C (**59-61**)

Despite their isolation in 2007, the synthesis of these molecules has not been achieved. Retrosynthetically, we hypothesize that these natural products could be accessed using chemistry developed for the synthesis of aspeverin (**1**). Spirocyclic pentacycle *ent*-PF1270A (***ent*-59**, drawn herein as its enantiomer for clarity) could be traced back via oxidative indole rearrangement to parent indole **62** (**Figure 5**). The prenyl epoxide moiety could be installed via metal-catalyzed cross-coupling of an indolic halide, as has previously been described by the Martin group for the synthesis of related alkaloids, citrinadins A and B.[33] In this case, one unique synthetic challenge would be installation of the quaternary stereocenter at position 3 with correct stereochemical configuration of the aldehyde and the methyl group. In the case of aspeverin, this quaternary carbon center was degenerate and took the form of a geminal dimethyl group. Presuming that our Me_3Al-mediated ring-opening/methylation sequence mechanistically delivers the methyl group from the top face of the molecule (see proposed intermediate **41**, **Scheme 10**), one could envision selective installation of this methyl group via ring opening of a similar compound such as **63**. Further disconnections involve installation of the carbamate

via Curtius rearrangement and a Diels−Alder/epimerization sequence, leading back to projected starting diene **65** and dienophile **66**.

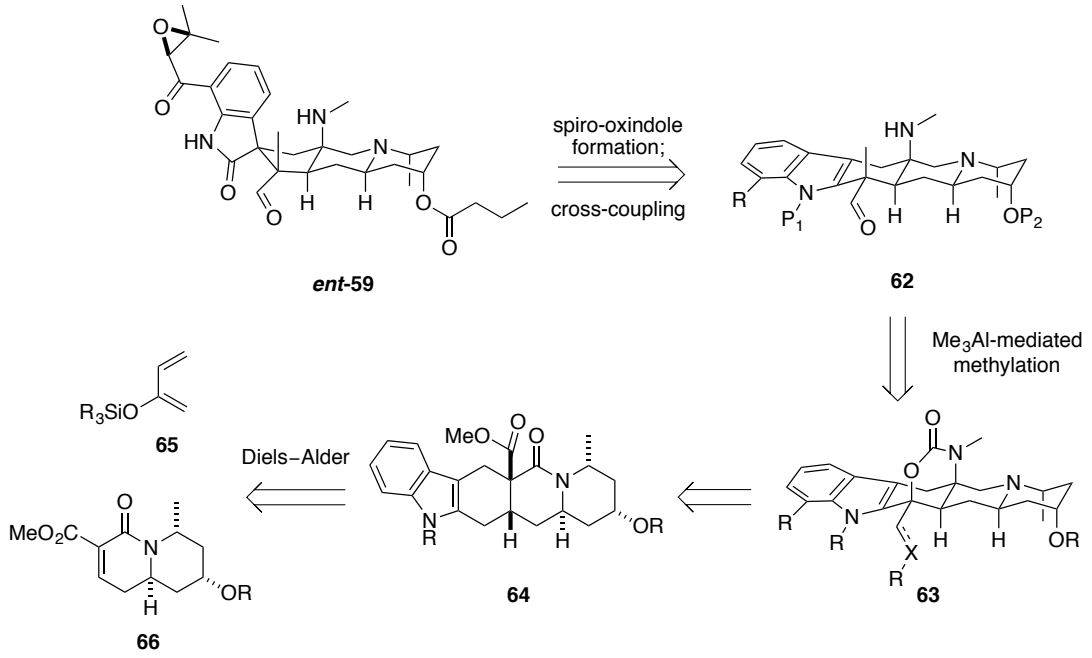
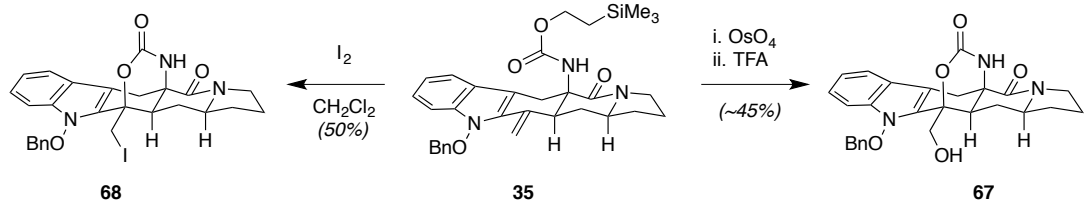


Figure 5: Retrosynthetic disconnections to access the PF1270 molecules (**59-61**) using chemistry developed for the synthesis of aspeverin (**1**).



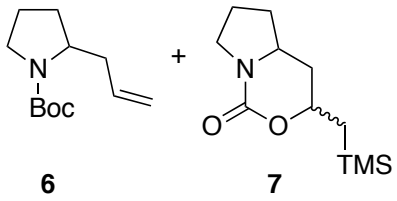
Scheme 12: divergent intramolecular cyclizations of carbamate **35** to access scaffolds similar to proposed intermediate **63**.

Preliminary model experiments utilizing an intermediate available from the aspeverin route found that we can indeed access compounds of type **63**. For example, dihydroxylation of exocyclic olefin **35** followed by treatment with acid affords the cyclic carbamate bearing an exocyclic hydroxymethylene in ca. 45% yield (**67**, **Scheme 12**). Moreover, this oxidation state can also be accessed via direct

iodocyclization of the carbamate to obtain exocyclic iodomethylene **68** in 50% isolated yield. Preliminary attempts at Me_3Al-mediated methylation using either of these compounds under similarly described conditions failed. The inductively electron withdrawing nature of the exocyclic heteroatoms (−OH or –I) may destabilize the presumed tertiary carbocation that was successfully ulitized in our synthesis of aspeverin. Alternatively, these heteroatoms may competitively coordinate to the Lewis acid used in this transformation. However, experimental efforts at this stage are not exhaustive. An in-depth analysis varying the nature of this exocyclic heteroatom or protection schemes of the hydroxymethylene as well as the nature of the Lewis acid and methyl nucleophile source may facilitate this desired transformation and open the door to the synthesis of related fungal natural products.

EXPERIMENTAL SECTION

General Information. All commercial reagents (Aldrich, Alfa-Aesar, Acros Organics, Fluka) were used without further purification. All solvents were reagent or HPLC grade (Fisher). Anhydrous solvents were purchased from Sigma-Aldrich and used without further purification. All reactions were carried out in flame-dried glassware under an argon or nitrogen atmosphere unless otherwise noted. Analytical TLC was performed on E. Merck silica gel 60 F254 plates and visualized by UV fluorescence quenching and $KMnO_4$ staining. Preparative thin-layer chromatography (PTLC) separations were performed on E. Merck silica gel 60 F254 plates (1 mm). Flash column chromatography was performed on Silicycle SiliaFlash F60 (40-63 μm). Yields refer to chromatographically and spectroscopically pure compounds. ^{1}H NMR and ^{13}C NMR spectra were recorded on a Bruker Avance DRX-500 MHz or DRX-600 MHz at ambient temperature unless otherwise stated. Chemical shifts are reported in parts per million relative to residual solvent $CDCl_3$ (^{1}H, 7.26 ppm, ^{13}C, 77.16 ppm), MeOH-d_4 (^{1}H, 3.31 ppm, ^{13}C, 49.0 ppm) or DMSO-d_6 (^{1}H, 2.50 ppm, ^{13}C, 39.52) ppm. Multiplicities are reported as follows: s = singlet, d = doublet, dd = doublet of doublets, t = triplet, td = triplet of doublets, tt = triplet of triplets, m = multiplet, q = quartet, app. = apparent, br. s = broad singlet. Diastereomeric ratio (dr) was determined by ^{1}H NMR analysis. High-resolution mass spectral analyses were performed by the MSKCC core facility staff.



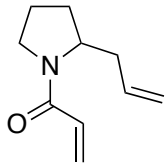
6 **7**

To a -78 ^{o}C solution of *tert*-butyl 2-oxopyrrolidine-1-carboxylate (**5**) (46.90 g, 253.2 mmol) in anhydrous Et_2O (135 mL) was added DIBAL-H (1.56 M in anhydrous PhMe, 187 mL, 291 mmol) by cannula over a period of ~30 minutes. The reaction was stirred for 3 hours at -78 ^{o}C, then 4 hours at room temperature. The reaction was recooled to -78 ^{o}C, and acidic MeOH (350 mg *p*TSA dissolved in 350 mL) was *carefully* added. The resulting mixture was stirred at room temperature overnight and concentrated. Saturated aq. sodium/potassium tartrate (400 mL) was added, and stirred for 3 hours until the mixture became turbid. The aqueous phase was extracted with ether (3 x 250 mL), the combined organic layers washed with brine (300 mL), dried over $MgSO_4$ and concentrated. The resulting crude hemiaminal was dissolved in CH_2Cl_2 (900 mL) and cooled in an ice bath. Allyltrimethylsilane (65 mL, 409 mmol) was added followed by $BF_3{\cdot}Et_2O$ (28 mL, 227 mmol) dropwise over 10 minutes. The reaction was stirred a further 15 minutes at 0 ^{o}C and 10 minutes at room temperature. Aqueous K_2CO_3 (600 mL) was added, and the organic layer was separated. The aqueous phase was further extracted with CH_2Cl_2 (2 x 300 mL), the combined organic layers washed with brine (500 mL), dried over $MgSO_4$ and concentrated. Purification by silica gel chromatography (10:1 Hexane/EtOAc → 1:1 Hexane/EtOAc→ 100% EtOAc) afforded **6** (30.0 g, 142 mmol, 56% yield) as a colorless oil and **7** (18.29 g, 32% yield, 1.2:1 *dr*) as a low-melting solid. Both **6** and **7** react productively in the next steps.

6 was in spectroscopic agreement as reported in the literature.[27]

^{1}H NMR (500 MHz, $CDCl_3$) mixture of rotamers δ 5.74 (m, 1 H), 5.04 (m, 2 H), 3.76-3.85 (m, 1 H), 3.31-3.38 (m, 2 H), 2.42-2.54 (m, 1 H), 2.11 (m, 1 H), 1.70-1.87 (m, 4 H), 1.46 (s, 9 H).

7: R_f=0.05 (silica gel, 2:1 Hexane/EtOAc). ^{1}H NMR (600 MHz, $CDCl_3$) δ 5.02 – 4.22 (m, 2H), 3.75 – 3.35 (m, 6H), 2.32 – 1.92 (m, 5H), 1.85– 1.69 (m, 3H), 1.54 – 1.27 (m, 4H), 1.27 – 1.10 (m, 2H), 0.95-0.80 (m, 2H), 0.08 (s, 18H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 153.5, 152.9, 76.3, 74.9, 56.8, 52.4, 46.7, 46.5, 36.8, 33.8, 33.5, 33.4, 24.6, 23.1, 23.0, 22.9, -0.7, -0.9. IR (film): cm^{-1} 2948, 2894, 1688, 1251, 843. HRMS (ESI, m/z) calcd for $C_{11}H_{22}NO_2Si$ $[M+H]^+$ 228.1420, found 228.1411.

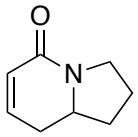


8

To a 0 ^{o}C solution of **6** (23.21 g, 109.8 mmol) in CH_2Cl_2 (75 mL) was added Et_3SiH (18 mL, 113 mmol) followed by careful addition of TFA (75 mL, 980 mmol) in small portions. The reaction was stirred for 1.5 hours, and concentrated to yield the corresponding amine TFA salt. The crude salt was then dissolved in CH_2Cl_2 (211 mL) and cooled to 0 ^{o}C. NEt_3 (60 mL, 430 mmol) was carefully added followed by 4-dimethylaminopyridine (670 mg, 5.5 mmol). Acryloyl chloride (11.5 mL, 142 mmol) was added dropwise over about 10 minutes, and the reaction was stirred for 2

hours. 1N HCl (200 mL) was then added, and the aqeous phase was extracted with CH_2Cl_2 (3 x 150 mL). The combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated. Silica gel chromatography (9:1 Hexane/EtOAc → 1:1 Hexane/EtOAc → 100% EtOAc) afforded **8** as a light yellow oil (13.51 g, 74% yield over 2 steps), which matched previously reported spectroscopic data.[2]

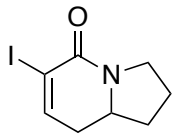
^{1}H NMR (500 MHz, $CDCl_3$) mixture of rotamers δ 6.40 (m, 2 H), 5.76 (m, 1 H), 5.68 (m, 1 H), 5.09 (m, 2 H), 4.00-4.26 (m, 1 H), 3.56 (m, 2 H), 2.32-2.65 (m, 2 H), 2.18 (m, 1 H), 1.76-2.02 (m, 4 H).



10

8 (11.72 g, 70.9 mmol) was dissolved in CH_2Cl_2 (1.4 L). The reaction was purged by bubbling N_2 through the solution for 20 minutes, at which point Grubbs' 2nd Generation catalyst (700 mg, .83 mmol, 1.17 mol%) was added. The reaction was stirred overnight (12 hours) under an atmosphere of N_2 at room temperature and was concentrated. Silica gel chromatography (1:1 CH_2Cl_2/Et_2O → 1:1 CH_2Cl_2/EtOAc → 20:1 CH_2Cl_2/MeOH) afforded **15** as a light brown oil (9.42 g, 96% yield). **15** matched previously reported spectroscopic data.[2]

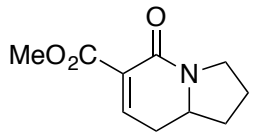
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.50 (m, 1 H), 5.96 (dd, *J* = 9.7, 2.9 Hz, 1 H), 3.67 (m, 2 H), 3.47 (m, 1 H), 2.47 (dt, *J* = 17.2, 5.6 Hz, 1 H), 2.22 (m, 1 H), 2.13 (m, 1 H), 2.02 (m, 1 H), 1.79 (m, 1 H), 1.61 (m, 1 H).



11

To a solution of **10** (9.42 g, 68.7 mmol) in CCl_4 (335 mL) and pyridine (335 mL) was added solid I_2 (53.0 g, 210 mmol). The resulting mixture was heated at 60 ^oC for 20 hours under an atmosphere of N_2. The reaction was cooled to room temperature and sat. aq. $Na_2S_2O_3$ (600 mL) was added followed by 1N HCl (200 mL). The organic layer was separated, and the aqueous phase extracted with EtOAc (3 x 300 mL). The combined organic layers were washed with brine (2 x 350 mL), dried over $MgSO_4$ and concentrated. Purification by silica gel chromatography (1:1 EtOAc/Hexane → 2:1 EtOAc/Hexane) yielded iodide **11** as a white solid (7.76 g, 43% yield). R_f = 0.59 (silica gel, 1:1 CH_2Cl_2/Et_2O)

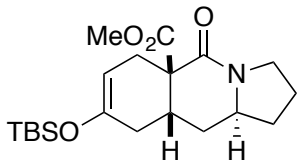
^1H NMR (500 MHz, $CDCl_3$) δ 7.18 (dd, *J* = 7.0, 2.3 Hz, 1H), 3.79 (m, 1H), 3.67 (m, 1H), 3.55 (m, 1H), 2.44 (ddd, *J* = 17.1, 6.9, 4.8 Hz, 1H), 2.33 – 2.15 (m, 2H), 2.10 (m, 1H), 1.81 (m, 1H), 1.60 (m, 1H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 159.1, 146.8, 98.2, 56.9, 45.8, 34.2, 33.6, 23.3. IR (film): cm^{-1} 3041, 2969, 2941, 2880, 1649, 1429, 1276. HRMS (ESI, m/z) calcd for $C_9H_{11}NOI$ $[M+H]^+$ 263.9885, found 263.9895.



12

Carbon monoxide gas was bubbled from a balloon through a solution of iodide **11** (1.80 g, 7.11 mmol) in MeOH (35 mL) and MeCN (180 mL) for 10 minutes. To the resulting solution was added $PdCl_2(dppf)\cdot CH_2Cl_2$ (500 mg, 0.61 mmol) followed by NEt_3 (6.50 mL, 46.6 mmol). Carbon monoxide was then bubbled through the solution for a further 10 minutes, and the mixture was then stirred under a balloon of CO gas at 60 ^oC for 1.5 hr. The reaction mixture was cooled to room temperature under air and concentrated under reduced pressure. The residue was washed with sat. aq. $NaHCO_3$ (80 mL) and extracted with CH_2Cl_2 (3 x 80 mL). The combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated. Purification by column chromatography (1:1 CH_2Cl_2/EtOAc→ 15:1 CH_2Cl_2/MeOH) yielded ester **16** as a light brown oil that solidifies upon standing (1.28 g, 93%). The product contains a contaminant with a mass of [M+CH_3OH], presumed as resulting from Michael addition of MeOH, and was carried forward in the next step and is presumed to react. Concentration of a sample from acetic acid gave a characterizable sample of **12**. R_f = 0.20 (silica gel, 1:1 CH_2Cl_2/Et_2O).

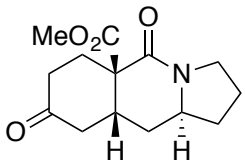
^1H NMR (600 MHz, $CDCl_3$) δ 7.34 (dd, *J* = 6.8, 2.2 Hz, 1H), 3.83 (s, 3H), 3.74 (m, 1H), 3.66 (m, 1H), 3.52 (ddd, *J* = 12.1, 10.1, 7.4 Hz, 1H), 2.61 (m, 1H), 2.35 – 2.19 (m, 2H), 2.04 (m, 1H), 1.82 (m, 1H), 1.67 (m, 1H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 165.2, 159.9, 145.6, 130.6, 56.2, 52.4, 44.7, 33.6, 30.9, 23.4. IR (film): cm^{-1} 2968, 2950, 2884, 1739, 1662, 1613, 1445, 1266. HRMS (ESI, m/z) calcd for $C_{10}H_{14}NO_3$ $[M+H]^+$ 196.0974, found 196.0973.



14

To a stirred solution of **12** (4.43 g, 22.73 mmol) in CH_2Cl_2 (225 mL) was added (buta-1,3-dien-2-yloxy)(*tert*-butyl)dimethylsilane (8.28 g, 45 mmol). Then $ZnCl_2$ (1M in Diethyl Ether, 23 mL, 23 mmol) was added in one portion and the reaction was stirred for 1.5 hours at room temperature. At this time, more diene was added (1.14 g, 6.2 mmol) and the reaction stirred a further 30 minutes until all starting material was consumed as determined by TLC analysis. The reaction was quenched by adding sat. aq. $NaHCO_3$ (250 mL). The layers were separated, and the aqueous layer was further extracted with CH_2Cl_2 (3 x 100 mL). The combined organic layers were washed with brine/sat. $NaHCO_3$ (1:1, 200 mL), dried over $MgSO_4$, and concentrated under reduced pressure. Silica gel chromatography (1:1 EtOAc/Hexanes → 2:1 EtOAc/Hexanes) afforded Diels−Alder adduct **14** (7.88 g, 91% yield) as a light orange oil and as a single diastereomer. R_f= 0.59 (silica gel, 1:1 CH_2Cl_2/Et_2O).

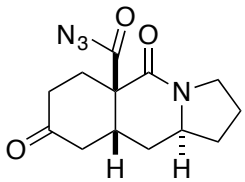
^{1}H NMR (500 MHz, $CDCl_3$) δ 4.84 (m, 1H), 3.77 (m, 1H), 3.73 (s, 3H), 3.59 (m, 1H), 3.30 (ddd, *J* = 12.7, 10.0, 3.0 Hz, 1H), 2.86 – 2.66 (m, 2H), 2.61 (ddt, *J* = 16.8, 4.2, 2.0 Hz, 1H), 2.19 (ddd, *J* = 17.7, 6.8, 2.0 Hz, 1H), 2.13 – 1.88 (m, 4H), 1.76 (m, 2H), 1.50 (m, 1H), 0.89 (s, 9H), 0.09 (s, 6H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 173.5, 168.3, 148.2, 101.3, 55.1, 53.6, 52.7, 44.6, 33.9, 32.9, 31.9, 29.2, 26.8, 25.8, 22.2, 18.1, -4.3, -4.4. IR (film): cm^{-1} 2953, 2934, 2892, 2856, 1730, 1647, 1442, 1247, 1178. HRMS (ESI, m/z) calcd for $C_{20}H_{34}NO_4Si$ $[M+H]^+$ 380.2257, found 380.2263.



15

To a stirred solution of cycloadduct **14** (62 mg, 0.163 mmol) in CH_2Cl_2 (2 mL) was added TFA (50 μL). The solution was stirred for 5 minutes and quenched with sat. aq. $NaHCO_3$ (4 mL). The mixture was diluted with CH_2Cl_2 (~8 mL) was added, and the mixture was directly filtered through a phase separator and concentrated. Purification by silica gel chromatography (1:1 EtOAc/CH_2Cl_2 → 10:1 CH_2Cl_2/MeOH) yielded pure ketone **15** as a colorless oil (43 mg, 99% yield). R_f = 0.22 (1:1 CH_2Cl_2/Et_2O)

^{1}H NMR (500 MHz, $CDCl_3$) δ 3.76 (m, 1H), 3.74 (s, 3H), 3.54 (m, 1H), 3.45 (ddd, *J* = 12.4, 9.7, 2.4 Hz, 1H), 2.75 (m, 1H), 2.54 (m, 1H), 2.50 – 2.28 (m, 5H), 2.10 (m, 2H), 2.05 (m, 1H), 1.90 (m, 2H), 1.84 (m, 1H), 1.52 (m, 1H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 210.5, 173.4, 166.6, 55.0, 53.4, 53.1, 45.4, 42.6, 37.2, 37.1, 33.6, 31.5, 29.4, 22.4. IR (film): cm^{-1} 2955, 2887, 1720, 1641, 1269, 1241, 1212. HRMS (ESI, m/z) calcd for $C_{14}H_{20}NO_4$ $[M+H]^+$ 266.1392, found 266.1380.

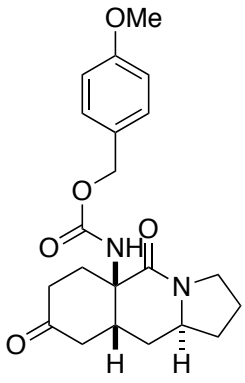


16

To a solution of **15** (1.27 g, 4.79 mmol) in MeOH (14 mL) and H_2O (4.8 mL) was added $LiOH \cdot H_2O$ (12.63 mmol). The reaction was left to stir overnight, and was

acidified with 1 M HCl. The aqueous mixture was extracted with EtOAc, organic layers were dried over $MgSO_4$, and concentrated to afford the corresponding crude acid (487 mg, ~42% yield). This crude material was suspended in toluene (10 mL), and cooled to 0 ^oC. To this solution was added triethylamine (0.45 mL, 3.23 mmol) followed by diphenylphosphoryl azide (0.49 mL, 2.28 mmol). The reaction was stirred for 1 hour, and was directly applied to a column of silica gel and the product was eluted (1:1 CH_2Cl_2/Et_2O) to afford acyl azide **16** (302 mg, 54%) as a colorless oil.

^1H NMR (500 MHz, $CDCl_3$) δ 3.74 (m, 1 H), 3.54 (m, 1 H), 3.47 (dt, *J* = 12.0, 1.7 Hz, 1 H), 2.77 (m, 1 H), 2.31-2.54 (m, 6 H), 2.12 (m, 1 H), 2.05 (m, 1 H), 1.79-1.86 (m, 3 H), 1.54 (m, 1 H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 210.0, 180.5, 166.0, 55.1, 54.7, 45.6, 42.6, 37.0, 36.8, 33.6, 31.6, 29.4, 22.5. IR (film): cm^{-1} 2964, 2884, 2141, 1708, 1631, 1453. LRMS: ESI(+)-MS for $C_{13}H_{16}N_4O_3$ *m/z:* 298.9 $[M+Na^+]^+$, 575.3 $[2M+Na^+]^+$.

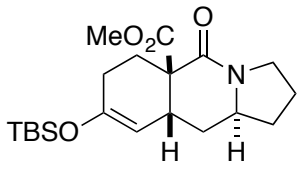


17

Acyl azide **16** (302 mg, 1.09 mmol) was dissolved in anhydrous toluene (10 mL) and heated to 85 ^oC. After 20 minutes, TLC analysis indicated full conversion to the

corresponding isocyanate. To the reaction was added 4-methoxybenzyl alcohol (1.0 mL, 8.06 mmol). The reaction was then sealed and heated in a 120°C oil bath for 5 hours. After cooling, the reaction was concentrated and purified by silica gel chromatography (1:1 CH_2Cl_2/Et_2O until excess PMBOH elutes, then flush with 1:1:0.25 CH_2Cl_2/Et_2O/MeOH) to afford carbamate **17** as a white solid (366 mg, 86% yield).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.27 (d, *J* = 8.6 Hz, 2 H), 6.87 (d, *J* = 8.6 Hz, 2 H), 5.49 (br s, 1 H), 5.02 (d, *J* = 11.8 Hz, 1 H), 4.96 (d, *J* = 11.8 Hz, 1 H), 3.80 (s, 3 H), 3.65 (m, 1 H), 3.50 (m, 1 H), 3.41 (m, 1 H), 3.27 (m, 1 H), 2.68 (dd, *J* = 15.1, 6.9 Hz, 1 H), 2.37-2.49 (m, 3 H), 2.17-2.31 (m, 2 H), 2.10 (m, 1 H), 1.97 (m, 1 H), 1.77-1.82 (m, 2 H), 1.50 (m, 1 H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 211.4, 170.0, 159.8, 156.1, 130.1, 128.4, 114.2, 66.8, 58.0, 55.5, 55.1, 46.2, 43.5, 35.7, 33.8, 33.4, 32.1, 22.9. LRMS: ESI(+)-MS for $C_{21}H_{26}N_2O_5$ *m/z:* 387.23 $[M+H^+]^+$, 773.73 $[2M+H^+]^+$.

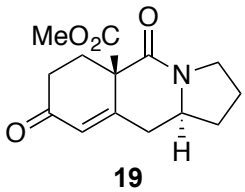


18

To a solution of **14** (692 mg, 1.83 mmol) in anhydrous toluene (62 mL) was added silica gel (40-63 μm, oven-dried for 24 hr, 1.89 g). The resulting mixture was stirred vigorously in a sealed vial at 110 °C 14 hours, at which point more silica gel (1.30 g) was added, and the mixture stirred a further 24 hours. During the course of the

reaction, small aliquots were periodically taken and reaction progress was monitored by ^{1}H NMR. The reaction was cooled and filtered to remove silica gel, which was washed with copious amounts of CH_2Cl_2. Purification by silica gel chromatography (1:1 CH_2Cl_2/Et_2O) yielded **18** as a colorless oil (469 mg, 10:1 ratio of product:starting material, 67% yield). R_f = 0.53 (1:1 CH_2Cl_2/Et_2O)

^{1}H NMR (600 MHz, $CDCl_3$) δ 4.66 (br s, 1H), 3.70 (s, 3H), 3.69 (m, 1H), 3.39 (m, 1H), 2.96 (m, 1H), 2.58 (m, 1H), 2.12 – 1.87 (m, 5H), 1.82 (m, 2H), 1.72 (m, 1H), 1.52 (m, 1H), 0.90 (s, 9H), 0.10 (s, 3H), 0.08 (s, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 174.4, 166.5, 153.0, 105.2, 55.8, 53.0, 52.8, 45.2, 37.1, 33.4, 31.8, 28.4, 26.7, 25.7, 22.5, 18.2, -4.1, -4.2. IR (film): cm^{-1} 2955, 2933, 2891, 2858, 1731, 1641, 1455, 1248, 1176, 840. HRMS (ESI, m/z) calcd for $C_{20}H_{34}NO_4Si$ $[M+H]^+$ 380.2257, found 380.2249.

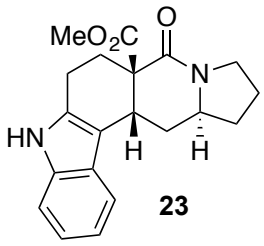


19

To a -78 °C stirred solution of **18** (469 mg, 1.23 mmol) in CH_2Cl_2 (22 mL) was added a solution of PhSeCl (345 mg, 1.81 mmol, dissolved in 2.70 mL CH_2Cl_2) dropwise over ca. 1 minute. The reaction was stirred for 10 minutes at -78 °C, then allowed to stir at room temperature for a further 2 minutes. The reaction was quenched with sat. aq. $NaHCO_3$ (50 mL), and the aqueous phase was extracted with CH_2Cl_2 (3 x 40 mL). The combined organic layers were washed with brine, dried

over $MgSO_4$ and concentrated. The crude mixture was quickly filtered through a short column of silica gel (1:1 CH_2Cl_2:Et_2O) to obtain the corresponding crude phenylselenide. This material was immediately dissolved in CH_2Cl_2 (13 mL), and H_2O_2 (0.4 mL, 30 wt% in H_2O) was added. The mixture was stirred for 15 minutes and quenched with sat. aq. $NaHCO_3$ (20 mL). The aqueous phase was extracted with CH_2Cl_2 (3 x 20 mL), and the combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated. Purification by silica gel chromatography (1:1 CH_2Cl_2/Et_2O) yielded pure enone **19** (168 mg, 51% yield overall) as a crystalline white solid. R_f = 0.35 (silica gel, 1:1 CH_2Cl_2/Et_2O). mp: 114-118°C

^{1}H NMR (600 MHz, $CDCl_3$) δ 5.96 (s, 1H), 3.75 (s, 3H), 3.66 (m, 1H), 3.52 (m, 1H), 3.45 (m, 1H), 3.03 (ddd, *J* = 13.8, 5.4, 2.0 Hz, 1H), 2.93 (ddd, *J* = 18.1, 13.8, 5.4 Hz, 1H), 2.83 (ddd, *J* = 12.8, 11.1, 1.7 Hz, 1H), 2.76 (dd, *J* = 12.8, 3.7 Hz, 1H), 2.51 (m 1H), 2.24 (m, 1H), 2.07 (m, 1H), 1.97 (td, *J* = 13.9, 5.0 Hz, 1H), 1.85 (m, 1H), 1.63 (m, 1H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 198.2, 168.7, 165.5, 156.0, 127.0, 57.7, 53.9, 53.4, 45.9, 38.0, 35.0, 33.8, 30.4, 22.4. IR (film): cm^{-1} 2977, 2958, 2926, 2887, 1739, 1653, 1628, 1441, 1344, 1223, 1166. HRMS (ESI, m/z) calcd for $C_{14}H_{18}NO_4$ $[M+H]^+$ 264.1236, found 264.1246.

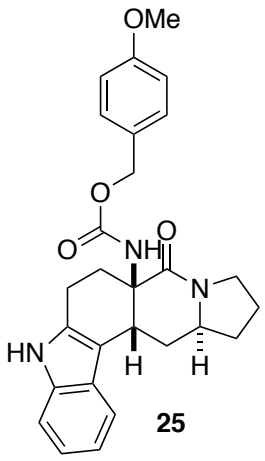


23

To a stirred solution of ketone **15** (111 mg, 0.42 mmol) in AcOH (0.84 mL) was added phenylhydrazine hydrochloride (79 mg, 0.55 mmol). The mixture was heated

at 90°C under N_2 for 3 hours, at which point the reaction was concentrated under reduced pressure. Purification by silica gel chromatography (1:1 CH_2Cl_2/Et_2O → 20:1 CH_2Cl_2/MeOH) afforded indole **23** as a light orange foam (96 mg, 67% yield). R_f = 0.22 (1:1 CH_2Cl_2/Et_2O)

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.94 (br s, 1H), 7.51 (d, *J* = 7.9 Hz, 1H), 7.32 (d, *J* = 7.9 Hz, 1H), 7.15 (t, *J* = 7.9 Hz, 1H), 7.08 (t, *J* = 7.9 Hz, 1H), 3.79 (s, 3H), 3.79 (m, 1H), 3.66 (dt, *J* = 12.8, 8.6 Hz, 1H), 3.34 (m, 1H), 2.98 (m, 1H), 2.86 (ddd, *J* = 12.5, 4.3, 2.3 Hz, 1H), 2.81 (dt, *J* = 13.2, 3.5 Hz, 1H), 2.75 – 2.60 (m, 2H), 2.23 (td, *J* = 12.5, 5.9 Hz, 1H), 2.08 (m, 1H), 1.90 (m, 2H), 1.55 (m, 2H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 174.4, 167.0, 136.2, 135.8, 126.4, 121.5, 119.6, 118.6, 111.0, 108.3, 55.3, 55.0, 52.9, 45.5, 36.5, 33.2, 30.3, 29.1, 22.3, 20.3. IR (film): cm^{-1} 3259, 2969m 2952, 2883, 1734, 1674, 1624, 1459, 1330, 1252. HRMS (ESI, m/z) calcd for $C_{20}H_{23}N_2O_3$ $[M+H]^+$ 339.1709, found 339.1715.

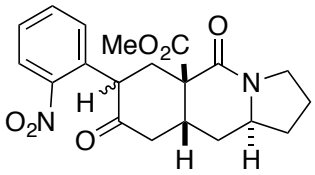


25

To a solution of ketone **17** (16.0 mg, 0.041 mmol) in AcOH (0.5 mL) was added phenylhydrazine hydrochloride (8.9 mg, 0.062 mmol). The reaction was then stirred at 80 °C for four hours. At this time, the reaction was cooled to room temperature

and concentrated. Purification by preparative TLC (1:1 CH_2Cl_2/Et_2O) afforded **25** as a colorless oil (6.7 mg, 36% yield)

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.86 (br s, 1 H), 7.51 (d, *J* = 7.8 Hz, 1 H), 7.29 (m, 3 H), 7.14 (t, *J* = 7.8 Hz, 1 H), 7.07 (t, *J* = 7.8 Hz, 1 H), 6.88 (d, *J* = 8.4 Hz, 2 H), 5.33 (br s, 1 H), 5.00 (m, 2 H), 4.41 (m, 1 H), 3.81 (s, 3 H), 3.58 (m, 1 H), 3.28 (m, 1 H), 3.00 (m, 1 H), 2.67 (m, 3 H), 2.52 (m, 2 H), 2.05 (m, 1 H), 1.88 (m, 2 H), 1.53 (m, 1 H), 1.46 (m, 1 H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 169.1, 159.8, 155.4, 136.8, 135.0, 130.1, 128.7, 127.0, 121.5, 119.6, 119.0, 114.1, 111.0, 109.7, 66.6, 60.7, 55.8, 55.5, 46.0, 35.5, 33.2, 31.3, 30.0, 22.5, 20.9.



27

A stirred solution of enol ether **14** (6.45 g, 17.01 mmol) in DMSO (34 mL) and THF (17 mL) was cooled in an ice bath for 10 minutes, until the mixture became a partially-frozen slurry. To the rapidly stirring mixture was added solid (*o*-nitrophenyl)phenyliodonium fluoride (7.33 g, 21.26 mmol) in one portion.[34] The resulting dark brown mixture was stirred for 2 minutes, at which time the ice bath was removed and the mixture stirred at room temperature for a further 20 minutes. Water (250 mL) was added, and the mixture was extracted with EtOAc (4 x 150 mL). The combined organic layers were washed with brine (100 mL), dried over $MgSO_4$ and concentrated. Silica gel chromatography (1:1 CH_2Cl_2:Et_2O) afforded a

2.2:1 diastereomeric mixture of nitroarene **27** (3.090 g, 47% yield) as a light yellow foam. For characterization purposes, the diastereomers can be easily separated by preparative TLC using a 1:1 CH_2Cl_2:Et_2O eluent system.

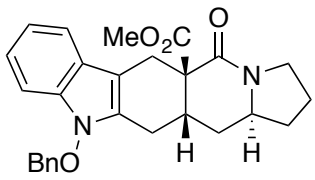
Nitroarene **S27**, major diastereomer: R_f = 0.50 (silica gel, 1:1 CH_2Cl_2/Et_2O)

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.08 (dd, *J* = 8.3, 1.4 Hz, 1H), 7.62 (td, *J* = 7.6, 1.4 Hz, 1H), 7.45 (ddd, *J* = 8.3, 7.6, 1.4 Hz, 1H), 7.39 (dd, *J* = 7.6, 1.4 Hz, 1H), 4.30 (m, 1H), 3.88 (m, 1H), 3.76 (s, 3H), 3.68 (m, 1H), 3.55 (ddd, *J* = 12.7, 9.9, 2.6 Hz, 1H), 3.26 (dd, *J* = 12.7, 4.7 Hz, 1H), 2.99 (m, 1H), 2.76 (t, *J* = 13.4 Hz, 1H), 2.70 – 2.54 (m, 2H), 2.21 – 2.03 (m, 2H), 2.03 – 1.88 (m, 2H), 1.83 (m, 1H), 1.54 (m, 1H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 206.1, 173.0, 165.4, 149.0, 133.8, 132.9, 129.8, 128.4, 125.6, 55.6, 55.5, 53.4, 49.4, 45.4, 43.0, 39.6, 35.7, 33.5, 30.4, 22.6. IR (film): cm^{-1} 2972, 2953, 2884, 1727, 1639, 1521, 1346, 908. HRMS (ESI, m/z) calcd for $C_{20}H_{23}N_2O_6$ $[M+H]^+$ 387.1556, found 387.1542.

Minor Diastereomer **27**: R_f = 0.43 (silica gel, 1:1 CH_2Cl_2/Et_2O)

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.03 (dd, *J* = 8.2, 1.3 Hz, 1H), 7.61 (td, *J* = 7.6, 1.3 Hz, 1H), 7.52 – 7.37 (m, 2H), 4.32 (br s, 1H), 3.81 (s, 3H), 3.67 (m, 2H), 3.50 (m, 1H), 3.10 – 2.87 (m, 2H), 2.73 (m, 1H), 2.64 (m, 1H), 2.36 (dd, *J* = 15.9, 3.4 Hz, 1H), 2.20 (dt, *J* = 11.9, 5.8 Hz, 1H), 2.12 (m, 1H), 2.09 – 1.97 (m, 2H), 1.87 (m, 1H), 1.61 (m, 1H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 207.67, 173.64, 167.66, 148.34, 134.13, 133.48, 128.66, 125.67, 54.21, 53.22, 52.25, 45.88, 42.58, 34.97, 34.22, 33.90, 31.84, 22.40. ^{13}C NMR (150 MHz, $CDCl_3$) δ 207.7, 173.7, 167.7, 148.4,

134.1, 133.5, 128.7, 125.7, 54.2, 53.2, 52.2, 45.9, 42.6, 35.0, 34.2, 33.9, 31.8, 22.4. IR (film): cm^{-1} 2953, 2884, 1727, 1635, 1530, 1350, 912. HRMS (ESI, m/z) calcd for $C_{20}H_{23}N_2O_6$ $[M+H]^+$ 387.1556, found 387.1541.

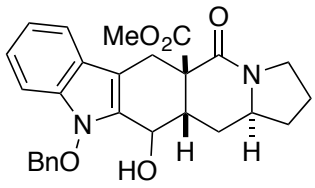


28

To a solution of nitroaryl ketone **27** (700 mg, 1.81 mmol) in EtOAc (24 mL) and AcOH (6 mL) was added 10% Pd/C (96 mg, 0.090 mmol) and $Pd(PPh_3)_4$ (32 mg, 0.028 mmol). H_2 was bubbled from a balloon through the reaction mixture for 10 minutes, and the reaction stirred under 1 atm H_2 for 8.5 hours. Reaction progress was monitored by occasionally withdrawing an aliquot and analyzing for disappearance of starting material by ^1H NMR in DMSO-*d*$_6$. The entire reaction contents were concentrated under reduced pressure, and the residue was resuspended in DMF (4.9 mL) and cooled to 0^oC. NaH (60% in mineral oil, 250 mg, 6.25 mmol) was then added, and the mixture stirred for 10 minutes. Benzyl bromide (0.35 mL, 2.95 mmol) was then added and the reaction was stirred at room temperature for 45 minutes until all starting material had disappeared by TLC analysis. The mixture was then recooled to 0^oC and carefully quenched with sat. aq. NH_4Cl (30 mL). The aqueous mixture was extracted with EtOAc (4 x 25 mL), the combined organic layers washed with brine (2 x 30 mL), dried over $MgSO_4$ and concentrated under reduced pressure. Purification by silica gel chromatography (1:1

CH_2Cl_2:Diethyl Ether) afforded alkylated *N*-hydroxyindole **28** as a yellow foam (758 mg, 94% yield). R_f = 0.60 (silica gel, 1:1 CH_2Cl_2/Et_2O).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.46 (d, *J* = 7.6 Hz, 1H), 7.43 – 7.30 (m, 6H), 7.18 (t, *J* = 7.6 Hz, 1H), 7.10 (t, *J* = 7.6 Hz, 1H), 5.13 (d, *J* = 10.9 Hz, 1H), 5.10 (d, *J* = 10.9 Hz, 1H), 3.76 (m, 1H), 3.71 (s, 3H), 3.55 – 3.43 (m, 3H), 3.32 (ddd, *J* = 12.7, 10.0, 2.9 Hz, 1H), 3.14 (d, *J* = 15.9 Hz, 1H), 2.90 (m, 1H), 2.49 (m, 1H), 2.28 (dd, *J* = 17.2, 4.8 Hz, 1H), 2.14 – 1.86 (m, 3H), 1.86 – 1.68 (m, 2H), 1.52 (m, 1H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.8, 168.5, 135.0, 133.8, 130.8, 130.2, 129.2, 128.8, 123.4, 121.9, 119.9, 118.5, 108.3, 103.2, 79.5, 55.0, 54.8, 52.8, 44.6, 34.1, 32.0, 28.7, 23.7, 23.5, 22.1. IR (film): cm^{-1} 3055, 3031, 2952, 2887, 1727, 1638, 1452, 1219, 904. HRMS (ESI, m/z) calcd for $C_{27}H_{29}N_2O_4$ $[M+H]^+$ 445.2127, found 445.2112.

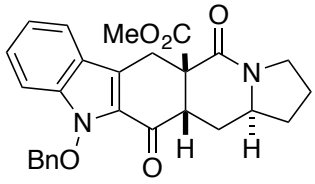


30

To a room temperature solution of **28** (1.470 g, 3.31 mmol) in CH_2Cl_2 (18 mL) was added $Pb(OAc)_4$ (1.830 g, 4.12 mmol) followed by acetic acid (9 mL). The resulting red solution was stirred at room temperature for 30 minutes and concentrated under reduced pressure. The residue was washed with 50 mL sat. aq. $NaHCO_3$ and extracted into CH_2Cl_2 (3 x 50 mL). The combined organic layers were washed with brine, dried over $MgSO_4$, filtered and concentrated. The crude acetate (**29**) was then dissolved in anhydrous MeOH (33 mL) and THF (10 mL). K_2CO_3 (1.830 g, 13.26

mmol) was added and the resulting mixture was stirred vigorously for 6.5 hours, at which point all of the acetate was consumed as judged by TLC. The mixture was quenched with sat. aq. NH_4Cl (50 mL), and extracted with EtOAc (4 x 50 mL). The combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated under reduced pressure. Silica gel chromatography (1:1 CH_2Cl_2/Diethyl Ether) afforded alcohol **30** as a light yellow foam and as a single unassigned diastereomer. (1.12 g, 73% overall). R_f = 0.34 (silica gel, 1:1 CH_2Cl_2/Et_2O).

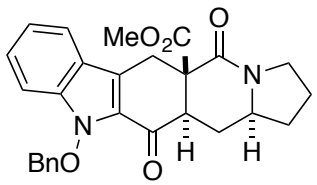
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.52 (d, *J* = 7.9 Hz, 1H), 7.47 – 7.32 (m, 6H), 7.24 (t, *J* = 7.4 Hz, 1H), 7.13 (t, *J* = 7.4 Hz, 1H), 5.33 (d, *J* = 10.4 Hz, 1H), 5.25 (d, *J* = 10.4 Hz, 1H), 4.42 (t, *J* = 5.0 Hz, 1H), 3.87 – 3.76 (m, 1H), 3.73 (s, 3H), 3.58 (d, *J* = 16.2 Hz, 1H), 3.45 (td, *J* = 12.4, 6.4 Hz, 1H), 3.34 (ddd, *J* = 12.4, 10.0, 3.0 Hz, 1H), 3.12 (d, *J* = 16.2 Hz, 1H), 3.02 (ddd, *J* = 10.0, 5.8, 3.8 Hz, 1H), 2.66 (d, *J* = 5.0 Hz, 1H), 2.15 – 1.94 (m, 2H), 1.88 – 1.70 (m, 1H), 1.54 (m, 1H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.7, 168.0, 134.9, 134.6, 131.6, 130.1, 129.3, 128.9, 123.3, 122.5, 120.3, 119.6, 108.8, 105.6, 80.0, 64.0, 55.2, 54.1, 52.9, 44.8, 41.0, 33.9, 26.8, 23.9, 22.1. IR (film): cm^{-1} 3368, 2953, 2888, 1723, 1616, 1465, 1320, 1213, 916. HRMS (ESI, m/z) calcd for $C_{27}H_{29}N_2O_5$ $[M+H]^+$ 461.2076, found 461.2064.



31

Alcohol **30** (1.331 g, 2.89 mmol) was dissolved in CH_2Cl_2 (120 mL), and activated MnO_2 (oven-dried for 12 hours, 13.30 g, 153 mmol) was added. The mixture was stirred at room temperature for 18 hours, at which point the reaction contents were filtered through a pad of celite. The filter cake was further washed with CH_2Cl_2 (~200 mL), and the filtrate was concentrated to yield ketone **31** (1.130g, 85% yield) as a pale yellow foam, which did not require any additional purification. R_f = 0.67 (1:1 CH_2Cl_2/Et_2O)

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.65 (d, *J* = 8.1 Hz, 1H), 7.56 (m, 2H), 7.48 – 7.36 (m, 3H), 7.33 (ddd, *J* = 8.2, 7.0, 1.1 Hz, 1H), 7.21 (d, *J* = 8.4 Hz, 1H), 7.12 (ddd, *J* = 8.1, 7.0, 1.1 Hz, 1H), 5.34 (d, *J* = 9.6 Hz, 1H), 5.25 (d, *J* = 9.6 Hz, 1H), 4.25 (d, *J* = 16.3 Hz, 1H), 3.83 (s, 3H), 3.67 (dt, *J* = 12.7, 8.8 Hz, 1H), 3.55 (m, 1H), 3.51 – 3.40 (m, 2H), 3.25 (ddd, *J* = 12.7, 9.8, 2.5 Hz, 1H), 2.99 (dt, *J* = 13.6, 4.3 Hz, 1H), 2.10 (m, 1H), 2.03 – 1.89 (m, 1H), 1.86 – 1.64 (m, 2H), 1.46 (m, 1H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 185.1, 172.8, 165.3, 137.7, 134.5, 130.2, 129.3, 128.8, 128.4, 127.4, 124.5, 121.8, 121.5, 120.8, 110.0, 80.8, 57.0, 56.2, 53.4, 48.4, 45.4, 33.4, 27.1, 26.8, 22.4. IR (film): cm^{-1} 3036, 3032, 2950, 2880, 1727, 1655, 1445, 1232, 1079. HRMS (ESI, m/z) calcd for $C_{27}H_{27}N_2O_5$ $[M+H]^+$ 459.1920, found 459.1904.

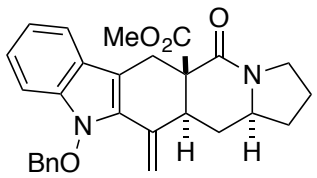


32

To a -70°C solution of KHMDS (1M in THF, 4.8 mL, 4.8 mmol) was added a solution of ketone **31** (1.130 g, 2.47 mmol) in THF (24 mL) dropwise over 10

minutes. The reaction was stirred for a further 5 minutes, and quenched with sat. aq. NH_4Cl (30 mL). The dry ice/acetone bath was removed, and 20 mL more sat. NH_4Cl was added. The mixture was extracted with EtOAc (4 x 50 mL), the combined organic layers washed with brine (2 x 50 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The compound was purified by silica gel chromatography (1:1 Hexane/EtOAc → 1:1 CH_2Cl_2/Et_2O) to afford *trans* ketone **32** (907 mg, 80% yield) as a light yellow solid. $R_f = 0.54$ (1:1 CH_2Cl_2/Et_2O)

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.64 (d, *J* = 7.8 Hz, 1H), 7.55 (m, 2H), 7.37 (m, 3H), 7.32 (t, *J* = 7.8 Hz, 1H), 7.25 (d, *J* = 7.5 Hz, 1H), 7.14 (t, *J* = 7.5 Hz, 1H), 5.33 (d, *J* = 9.6 Hz, 1H), 5.24 (d, *J* = 9.6 Hz, 1H), 4.24 (d, *J* = 16.6 Hz, 1H), 3.69 (dt, *J* = 12.6, 8.8 Hz, 1H), 3.54 (s, 3H), 3.52 (m, 1H), 3.13 (d, *J* = 16.6 Hz, 1H), 3.03 (dd, *J* = 12.1, 2.6 Hz, 1H), 2.86 (ddd, *J* = 13.6, 4.4, 2.8 Hz, 1H), 2.31 – 2.09 (m, 2H), 2.04 (m, 1H), 1.87 (m, 1H), 1.62 (m, 1H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 185.9, 170.3, 164.7, 136.9, 134.5, 130.3, 129.3, 128.7, 127.8, 127.5, 121.8, 121.4, 121.3, 120.8, 110.1, 80.8, 59.5, 57.1, 53.3, 51.9, 45.5, 33.2, 30.9, 25.2, 22.2. IR (film): cm^{-1} 3069, 3031, 2973, 2952, 2872, 1724, 1681, 1638, 1438, 1205, 1176. HRMS (ESI, m/z) calcd for $C_{27}H_{27}N_2O_5$ $[M+H]^+$ 459.1920, found 459.1908.

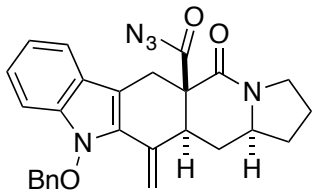


34

To a 0°C solution of ketone **32** (855 mg, 1.86 mmol) in THF (25 mL) was added a solution of Tebbe Reagent (0.5M in PhMe, 13 mL, 6.5 mmol) over a period of about

1 minute. The resulting solution was stirred at room temperature for 30 minutes, recooled to 0°C and carefully quenched by the slow addition of sat. aq. $NaHCO_3$ (60 mL). The mixture was extracted with EtOAc (4 x 50 mL), the combined organic layers washed with brine (2 x 50 mL), dried over $MgSO_4$ and concentrated. Purification by silica gel chromatography (1:1 CH_2Cl_2:Et_2O) yielded olefin **34** as a light yellow foamy solid (657 mg, 77% yield). R_f = 0.54 (1:1 CH_2Cl_2/Et_2O)

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.58 (d, *J* = 7.9 Hz, 1H), 7.47 – 7.34 (m, 7H), 7.26 (t, *J* = 7.9 Hz, 1H), 7.14 (t, *J* = 7.5 Hz, 1H), 5.97 (s, 1H), 5.16 (s, 1H), 4.94 (d, *J* = 9.7 Hz, 1H), 4.90 (d, *J* = 9.7 Hz, 1H), 4.01 (d, *J* = 16.1 Hz, 1H), 3.72 (m, 1H), 3.61 (m, 1H), 3.55 (m, 1H), 3.48 (s, 3H), 3.00 (m, 1H), 2.87 (d, *J* = 16.1 Hz, 1H), 2.54 – 2.31 (m, 2H), 2.23 (m, 1H), 2.06 (m, 1H), 1.87 (m, 1H), 1.70 (m, 1H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 170.7, 166.9, 135.4, 135.3, 134.3, 130.6, 129.7, 129.2, 128.8, 123.8, 123.2, 120.6, 119.7, 110.9, 109.2, 107.8, 78.3, 59.1, 55.4, 52.6, 45.6, 44.9, 33.6, 29.9, 27.2, 22.4. IR (film): cm^{-1} 3066, 3031, 2973, 2955, 2883, 1728, 1680, 1638, 1330, 1205, 735. HRMS (ESI, m/z) calcd for $C_{28}H_{29}N_2O_4$ $[M+H]^+$ 457.2127, found 457.2107.

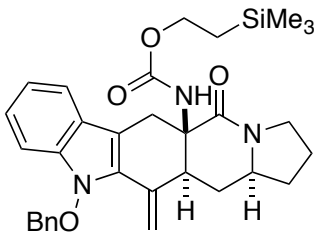


S-1

Ester **34** (500 mg, 1.09 mmol) was dissolved in 1,4-Dioxane (10 mL) and MeOH (10 mL). To the resulting solution was added aq. NaOH (820 mg dissolved in 1.25 mL H_2O, 20.5 mmol), and the mixture was heated at 55°C for 18 hours. The

reaction mixture was cooled to room temperature and quenched by the addition of 0.5M HCl (75 mL). The mixture was extracted with EtOAc (4 x 50 mL), the combined organic layers washed with brine (2 x 50 mL), dried over $MgSO_4$ and concentrated. The crude carboxylic acid was immediately resuspended in anhydrous toluene (11 mL), and NEt_3 (0.95 mL) and diphenylphosphoryl azide (0.95 mL) were added. The mixture was stirred for 45 minutes at room temperature, and the entire reaction contents were loaded onto a column of silica gel and purified (1:1 EtOAc/Hexanes → 2:1 EtOAc/Hexanes) to yield acyl azide **S-1** as a colorless foam (362 mg, 71% yield overall). R_f = 0.70 (silica gel, 1:1 CH_2Cl_2/Et_2O).

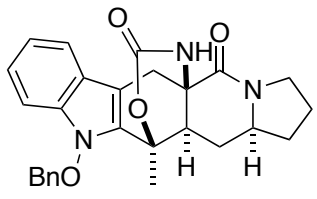
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.58 (d, *J* = 7.8 Hz, 1H), 7.40 (m, 6H), 7.29 (d, *J* = 7.2 Hz, 1H), 7.16 (t, *J* = 7.8 Hz, 1H), 6.03 (s, 1H), 5.22 (s, 1H), 4.98 (d, *J* = 9.7 Hz, 1H), 4.93 (d, *J* = 9.7 Hz, 1H), 3.98 (d, *J* = 16.5 Hz, 1H), 3.71 (m, 1H), 3.64 – 3.48 (m, 2H), 2.99 (dd, *J* = 11.8, 3.0 Hz, 1H), 2.90 (d, *J* = 16.5 Hz, 1H), 2.58 – 2.36 (m, 2H), 2.24 (m, 1H), 2.09 (m, 1H), 1.90 (m, 1H), 1.72 (m, 1H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 178.3, 165.9, 135.2, 134.6, 134.1, 130.5, 129.6, 129.2, 128.8, 124.0, 122.9, 120.7, 119.6, 110.0, 109.1, 108.4, 78.4, 59.1, 56.3, 45.7, 44.9, 33.4, 30.0, 27.0, 22.3. IR (film): cm^{-1} 3062, 3037, 2973, 2943, 2891, 2128, 1716, 1628, 1173. HRMS (ESI, m/z) calcd for $C_{27}H_{26}N_5O_3$ $[M+H]^+$ 468.2036, found 468.2029.



35

Acyl azide **S-1** (362 mg, 0.77 mmol) was dissolved in anhydrous toluene (5.6 mL) and heated to 90°C for 20 minutes under a nitrogen atmosphere (N_2 starts visibly bubbling from reaction mixture within about 1 minute of heating at 90°C). To the resulting formed isocyanate was added 2-(trimethylsilyl)ethanol (1.4 mL, 9.76 mmol) and the reaction mixture was heated at 90 °C for 9 hours, at which point the entire reaction contents were loaded onto a column of silica gel and eluted (2:1 Hexanes/Acetone → 1:1 Hexanes/Acetone) to yield carbamate **35** as a light yellow foamy solid (285 mg, 66% yield). R_f = 0.62 (silica gel, 1:1 CH_2Cl_2/Et_2O)

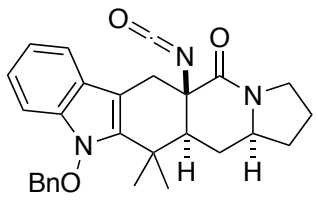
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.51 (d, *J* = 7.8 Hz, 1H), 7.48 – 7.36 (m, 6H), 7.29 (t, *J* = 7.6 Hz, 1H), 7.15 (t, *J* = 7.8 Hz, 1H), 6.13 (s, 1H), 5.26 (s, 1H), 5.05 (d, *J* = 9.6 Hz, 1H), 4.98 (d, *J* = 9.6 Hz, 1H), 4.86 (s, 1H), 3.98 (m, 2H), 3.84 (m, 1H), 3.75 – 3.57 (m, 1H), 3.52 (m, 2H), 3.07 (d, *J* = 16.5 Hz, 1H), 2.94 (m, 1H), 2.26 (m, 1H), 2.19 – 1.98 (m, 2H), 1.82 (m, 1H), 0.82 (m, 2H), -0.05 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 167.9, 155.9, 135.0, 134.6, 134.0, 130.2, 129.8, 129.4, 128.9, 124.1, 123.1, 120.8, 119.3, 110.2, 109.1, 106.4, 78.9, 63.1, 59.7, 57.6, 47.2, 45.7, 34.5, 32.8, 25.7, 22.9, 17.5, -1.4. IR (film): cm^{-1} 3431, 3062, 2952, 2893, 1724, 1641, 1330, 1247, 1226. HRMS (ESI, m/z) calcd for $C_{32}H_{40}N_3O_4Si$ $[M+H]^+$ 558.2788, found 558.2787.



40

To a room temperature solution of carbamate **35** (200 mg, 0.34 mmol) in CH_2Cl_2 (18 mL) was added TFA (50 mL, 0.65 mmol), and the mixture was stirred for 30 minutes, whereupon more TFA (50 μL, 0.65 mmol) was added. After stirring for an additional 45 minutes, sat. aq. $NaHCO_3$ (20 mL) was added. The aqueous phase was further extracted with CH_2Cl_2 (3 x 30 mL), the combined organic layers washed with brine, dried over $MgSO_4$ and concentrated. Purification by silica gel chromatography (1:1 CH_2Cl_2/Et_2O → 1:1:0.1 CH_2Cl_2/Et_2O/MeOH) afforded bicyclic carbamate **40** as a white crystalline solid (124 mg, 80% yield). R_f = 0.22 (1:1 CH_2Cl_2/Et_2O).

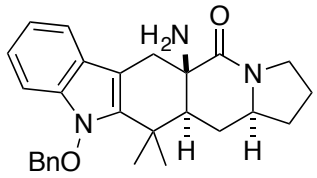
^{1}H NMR (600 MHz, $CDCl_3$) δ 7.56 (m, 2H), 7.52 (d, *J* = 7.6 Hz, 1H), 7.48 – 7.39 (m, 4H), 7.30 (t, *J* = 7.6 Hz, 1H), 7.16 (t, *J* = 7.6 Hz, 1H), 5.48 (s, 1H), 5.22 (d, *J* = 9.4 Hz, 1H), 5.08 (d, *J* = 9.4 Hz, 1H), 3.67 (m, 1H), 3.63 – 3.45 (m, 3H), 2.85 (d, *J* = 16.3 Hz, 1H), 2.46 (m, 1H), 2.34 (m, 1H), 2.26 (m, 1H), 2.10 (m, 4H), 1.88 (m, 1H), 1.61 (m, 1H), 1.49 (m, 1H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 167.7, 152.4, 134.7, 134.2, 130.7, 130.0, 129.3, 128.9, 124.4, 122.2, 121.0, 119.5, 109.6, 107.5, 80.2, 75.5, 58.5, 56.8, 45.7, 43.1, 33.8, 33.6, 25.3, 22.6, 20.2. IR (film): cm^{-1}3242, 3054, 2962, 2887, 1702, 1638, 1344, 1068. HRMS (ESI, m/z) calcd for $C_{27}H_{28}N_3O_4$ $[M+H]^+$ 458.2065, found 458.2080.



43

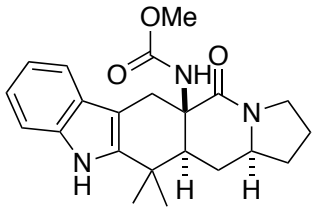
To a 0°C solution of **40** (145 mg, 0.317 mmol) in CH_2Cl_2 (11.5 mL) was added Me_3Al dropwise (2M solution in toluene, 2.90 mL, 5.80 mmol). The reaction was stirred for 1.5 hours at 0°C, and was carefully quenched with sat. aq. $NaHCO_3$ (30 mL). The mixture was extracted with EtOAc (4 x 40 mL), the combined organic layers washed with brine, dried over $MgSO_4$ and concentrated. The crude amine (**42**) was then redissolved in CH_2Cl_2 (3 mL), and 4-dimethylaminopyridine (48 mg, 0.39 mmol) was added followed by Boc_2O (0.10 mL, 0.43 mmol). The reaction was stirred for 1 hour, and loaded directly onto preparative TLC plates (100% Hexanes then 1:1 CH_2Cl_2/Et_2O) to yield isocyanate **43** (49 mg, 34% yield) as an amorphous white solid. $R_f = 0.80$ (silica gel, 1:1 CH_2Cl_2/Et_2O)

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.65 – 7.41 (m, 6H), 7.25 (d, *J* = 7.4 Hz, 3H), 7.14 (t, *J* = 7.4 Hz, 1H), 5.33 (d, *J* = 10.1 Hz, 1H), 5.23 (d, *J* = 10.1 Hz, 1H), 3.78 (dt, *J* = 12.2, 8.9 Hz, 1H), 3.62 (d, *J* = 16.1 Hz, 1H), 3.52 – 3.31 (m, 2H), 2.92 (d, *J* = 16.1 Hz, 1H), 2.28 – 2.14 (m, 2H), 2.07 (m, 1H), 1.92 (dd, *J* = 12.3, 1.8 Hz, 1H), 1.89 – 1.80 (m, 1H), 1.60 (m, *J* = 4.1 Hz, 1H), 1.55 (s, 3H), 1.54 (s, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 168.3, 137.9, 134.6, 133.8, 129.1, 129.0, 128.8, 126.8, 123.4, 122.4, 120.2, 118.8, 108.7, 102.3, 78.6, 61.8, 59.9, 49.5, 45.5, 35.6, 34.2, 33.5, 28.6, 26.1, 22.4, 21.4. IR (film): cm^{-1} 2969, 2894, 2234 (sharp), 1649, 1452. HRMS (ESI, m/z) calcd for $C_{28}H_{30}N_3O_3$ $[M+H]^+$ 456.2287, found 456.2280.



42

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.54 (m, 3 H), 7.40-7.48 (m, 4 H), 7.23 (ddd, *J* = 6.7, 6.0, 0.9 Hz, 1 H), 7.12 (ddd, *J* = 6.7, 6.0, 0.8 Hz, 1 H), 5.31 (d, *J* = 8.4 Hz, 1 H), 5.26 (d, *J* = 8.4 Hz, 1 H), 3.71 (m, 1 H), 3.47 (m, 2 H), 3.34 (d, *J* = 13.0 Hz, 1 H), 2.80 (d, *J* = 13.0 Hz, 1 H), 2.17 (m, 1 H), 2.11 (m, 1 H), 2.04 (m, 1 H), 1.97 (dd, *J* = 10.3, 1.1 Hz, 1 H), 1.84 (m, 1 H), 1.77 (m, 1 H), 1.60 (m, 1 H), 1.55 (s, 3 H), 1.54 (s, 3 H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 173.2, 138.2, 134.7, 134.1, 129.2, 129.0, 128.9, 124.1, 122.4, 120.1, 119.0, 108.7, 103.4, 78.6, 60.1, 56.8, 49.8, 45.3, 35.4, 34.9, 33.6, 29.2, 25.6, 22.6, 22.1. UPLC/MS: ESI(+)-MS for $C_{27}H_{31}N_3O_2$ *m/z*: 430.4 $[M+H^+]^+$.

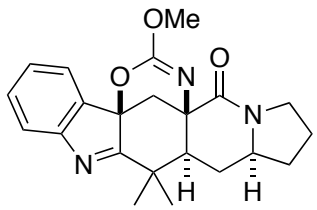


44

Isocyanate **43** (61 mg, 0.13 mmol) was suspended in 5 mL anhydrous MeOH and heated in a Biotage microwave reactor (closed system) at 100 ^{o}C for 13 hours. The resulting solution was concentrated and purified by preparative TLC (1:1 CH_2Cl_2/Et_2O) to yield carbamate **44** as an amorphous white solid (43 mg, 84%). R_f= 0.26 (1:1 CH_2Cl_2/Et_2O)

^{1}H NMR (500 MHz, $CDCl_3$): δ 8.15 (s, 1H), 7.48 (d, *J* = 7.6 Hz, 1H), 7.35 (d, *J* = 7.6 Hz, 1H), 7.18 (t, *J* = 7.6 Hz, 1H), 7.12 (t, *J* = 7.6 Hz, 1H), 4.92 (s, 1H), 3.95 – 3.78 (m, 1H), 3.61 – 3.34 (m, 6H), 2.92 (d, *J* = 15.8 Hz, 1H), 2.25 – 1.99 (m, 6H), 1.97 – 1.70 (m, 2H), 1.40 (s, 3H), 1.34 (s, 3H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 168.8, 156.2, 141.1, 136.6, 127.7, 122.0, 119.8, 118.5, 110.9, 104.2, 60.5, 57.4,

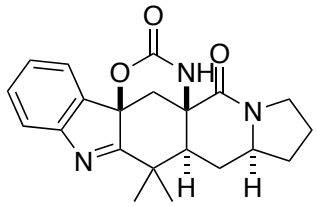
52.1, 50.3, 45.7, 34.8, 34.3, 32.7, 30.3, 24.7, 22.9. IR (film): cm^{-1} 3424, 3266, 2969, 2869, 1724, 1634, 1502, 1455, 1244, 911. HRMS (ESI, m/z) calcd for $C_{22}H_{27}N_3O_3$ $[M+H]^+$ 382.2131, found 382.2137



56

To a solution of **44** (10.0 mg, 0.026 mmol) in 1,1,1,3,3,3-hexafluoro-2-propanol (0.8 mL) was added $PhI(OAc)_2$ (14.0 mg, 0.043 mmol). The reaction was stirred for 30 minutes and was then concentrated. Purification by preparative TLC afforded **56** as a colorless oil (5.5 mg, 55% yield).

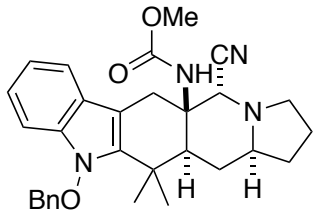
^1H NMR (600 MHz, $CDCl_3$) δ 7.55 (d, *J* = 7.6 Hz, 1 H), 7.45 (d, *J* = 7.2 Hz, 1 H), 7.40 (m, 1 H), 7.26 (m, 1 H), 3.78 (m, 1 H), 3.77 (s, 3 H), 3.47 (m, 1 H), 3.41 (m, 1 H), 3.11 (d, *J* = 14.0 Hz, 1 H), 2.15 (m, 1 H), 2.06 (m, 3 H), 2.00 (m, 1 H), 1.81-1.69 (m, 2 H), 1.65 (m, 1 H), 1.50 (d, *J* = 14.0 Hz, 1 H), 1.46 (s, 3 H), 1.14 (s, 3 H). ^{13}C NMR (150 MHz, $CDCl_3$): δ 186.7, 169.5, 153.4, 153.3, 136.2, 130.8, 126.9, 123.6, 131.1, 86.5, 60.6, 58.1, 55.5, 53.4, 45.0, 39.8, 35.5, 33.3, 28.0, 25.6, 22.5, 21.4. LRMS: ESI(+)-MS for $C_{22}H_{25}N_3O_3$ *m/z*: 380.9 $[M+H^+]^+$, 402.8 $[M+Na^+]^+$, 782.3 $[2M+Na^+]^+$.



penicimutamide A (**2**)

56 (5.4 mg, 0.0142 mmol) was massed into a 1-dram glass vial, and a solution of sodium 2-methyl-2-propanethiolate (23.9 mg, 0.21 mmol) in anhydrous DMF (0.45 mL) was added. The vial was sealed under argon and stirred in a 60 ^{o}C oil bath for 55 minutes, at which point UPLC/MS analysis indicated full conversion. The reaction was cooled to room temperature and quenched upon the addition of saturated aqueous NH_4Cl (2 mL) and brine (3 mL). The mixture was extracted with CH_2Cl_2 (3 x 8 mL). The combined organic layers were further washed with brine, filtered through a phase separator and concentrated. Purification by preparative TLC (20:1 CH_2Cl_2/MeOH) afforded **2** as a light yellow solid (4.1 mg, 79% yield).

^{1}H NMR (600 MHz, MeOH-d_4) δ 7.56 (d, J = 7.6 Hz, 1 H), 7.52 (d, J = 7.6 Hz, 1 H), 7.47 (td, J = 7.6, 1.0 Hz, 1 H), 7.34 (td, J = 7.6, 1.0 Hz, 1 H), 3.62 (td, J = 12.3, 9.0 Hz, 1 H), 3.53 (m, 1 H), 3.39 (ddd, J = 12.2, 10.0, 1.7 Hz, 1 H), 3.17 (d, J = 14.0 Hz, 1 H), 2.23 (ddd, J = 13.0, 4.0, 1.6 Hz, 1 H), 2.17 (m, 1 H), 2.05 (m, 1 H), 2.04 (dd, J = 12.6, 1.6 Hz, 1 H), 1.87 (m, 2 H), 1.73 (td, J = 13.0, 11.5 Hz, 1 H), 1.65 (m, 1 H), 1.50 (s, 3 H), 1.28 (s, 3 H). ^{13}C NMR (150 MHz, MeOH-d_4): δ 187.5, 167.4, 155.2, 153.8, 137.3, 132.0, 128.2, 124.4, 121.7, 88.0, 61.6, 58.2, 52.3, 46.6, 40.6, 36.9, 33.5, 27.7, 25.7, 23.2, 21.6. LRMS: ESI(+)-MS for $C_{21}H_{23}N_3O_3$ *m/z*: 366.2 $[M+H^+]^+$, 388.2 $[M+Na^+]^+$

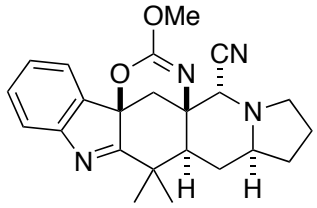


58

Lactam **44** (43 mg, 0.11 mmol) was dissolved in THF (1.14 mL) in a flame dried 25 mL round bottom flask. The solution was cooled to -45^{o}C in an acetone/dry ice bath, and DIBAL-H (1M in PhMe, 1.0 mL, 1.0 mmol) was added dropwise over 1 minute. The reaction was then stirred for 2 hours, maintaining the reaction bath temperature between -40 ^{o}C and -30 ^{o}C the entire time by periodically adding more dry ice. The reaction was then allowed to warm to -15 ^{o}C over 20 minutes, and was quenched with an aqueous solution of KCN (375 mg dissolved in 1.18 mL H_2O). The mixture was stirred for 45 minutes vigorously at room temperature, at which point sat. aq. Na/K tartrate (8 mL) and sat. $NaHCO_3$ (2 mL) were added and stirred vigorously until the reaction mixture became turbid. The mixture was further diluted with 8 mL sat. aq. Na/K tartrate and extracted with EtOAc (3 x 20 mL), the combined organic layers were washed with brine (15 mL), dried over $MgSO_4$ and concentrated . Purification by preparative TLC (1:1 Hexane/EtOAc) afforded **58** as an amorphous white solid and as a single diastereomer (21 mg, 47% yield). Further elution of the preparative TLC plate (1:1 CH_2Cl_2/EtOAc) afforded recovered starting material **44** (7.5 mg, 17% yield, 64% brsm). R_f= 0.53 (silica gel, 1:1 Hexane/EtOAc).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79 (br s, 1H), 7.51 (d, *J* = 7.8 Hz, 1H), 7.31 (d, *J* = 7.8 Hz, 1H), 7.15 (t, *J* = 7.4 Hz, 1H), 7.10 (t, *J* = 7.4 Hz, 1H), 5.06 (s, 1H), 4.78 (s,

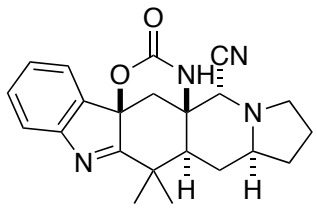
1H), 4.03 (d, *J* = 16.0 Hz, 1H), 3.51 (s, 3H), 3.04 (d, *J* = 16.0 Hz, 1H), 3.00 (dt, *J* = 8.6, 2.7 Hz, 1H), 2.54 (m, 2H), 2.09 (dd, *J* = 13.0, 3.5 Hz, 1H), 2.01 (m, 2H), 1.89 (m, 1H, 1.81 (m, 1H), 1.52 (m, 1H), 1.39 (s, 3H), 1.32 (s, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 155.2, 139.0, 136.4, 127.6, 121.9, 119.6, 118.7, 115.9, 110.6, 105.1, 58.4, 58.2, 56.4, 51.9, 50.9, 46.6, 34.0, 30.7, 30.5, 27.4, 24.9, 21.3. IR (film): cm^{-1} 3370, 2962, 2880, 2822, 1713, 1512, 1463, 1241, 907. HRMS (ESI, m/z) calcd for $C_{23}H_{28}N_4O_5$ $[M+H]^+$ 393.2291, found 393.2300.



S-2

To a room temperature solution of **58** (21.0 mg, 0.053 mmol) in 1,1,3,3-hexafluoro-2-propanol (2 mL) was added $PhI(OAc)_2$ (23 mg, 0.071 mmol), and the mixture was allowed to stand for 8 minutes. The reaction mixture was concentrated under reduced pressure and purified by preparative TLC (1:1 Hexane/EtOAc) to yield **S-2** (15.0 mg, 71% yield) as a colorless oil. R_f= 0.50 (silica gel, 1:1 Hexane/EtOAc). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.57 (d, *J* = 7.7 Hz, 1H), 7.43 (t, *J* = 7.7 Hz, 1H), 7.39 (d, *J* = 7.4 Hz, 1H), 7.26 (t, 7.4 Hz, 1 H), 3.83 (s, 1H), 3.80 (s, 3H), 3.05 (td, *J* = 8.6, 2.9 Hz, 1H), 2.54 (q, *J* = 8.6 Hz, 1H), 2.42 (m, 1H), 2.19 (d, *J* = 13.4 Hz, 1H), 2.04 – 1.94 (m, 2H), 1.91 (d, *J* = 13.4 Hz, 1H), 1.90 – 1.73 (m, 3H), 1.68 (m, 1H), 1.59 (m, 1H), 1.45 (s, 3H), 1.09 (s, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 186.2, 153.4, 152.8, 135.6, 131.0, 126.7, 123.0, 121.3, 115.3, 86.0, 63.4, 59.3, 55.9, 55.6,

51.7, 50.8, 39.9, 38.3, 30.3, 27.5, 27.4, 21.5, 21.4. IR (film): cm^{-1} 2969, 2955, 2869, 2819, 1674, 1580, 1308, 1262, 904. HRMS (ESI, m/z) calcd for $C_{23}H_{26}N_4O_2$ $[M+H]^+$ 391.2134, found 391.2146.



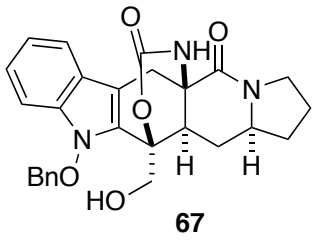
aspeverin (**1**)

A solution of sodium 2-methyl-2-propanethiolate (50 mg, 0.45 mmol) in anhydrous DMF (1.2 mL) was added to **S-2** (15.0 mg, 0.038 mmol) in a 10 mL round bottom flask. The stirred mixture was heated to 60°C for 1.5 hours, at which point the reaction was cooled to room temperature and quenched with sat. aq. NH_4Cl (5 mL) and brine (5 mL). The mixture was extracted with EtOAc (3 x 10 mL), the combined organic layers washed with brine (5 mL), dried over $MgSO_4$ and concentrated. Purification by preparative TLC (1:1 Hexane/EtOAc) afforded **1** as a colorless gum (11.1 mg, 74% yield). R_f = 0.38 (silica gel, 1:1 Hexane/EtOAc)

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.58 (dd, *J* = 8.1, 1.0 Hz, 1H), 7.50 – 7.35 (m, 2H), 7.28 (td, *J* = 7.5, 1.0 Hz, 1H), 6.25 (s, 1H), 3.92 (s, 1H), 3.00 (td, *J* = 8.8, 2.7 Hz, 1H), 2.55 (q, *J* = 8.8 Hz, 2H), 2.50 (m, 1H), 2.47 (d, *J* = 13.5 Hz, 1H), 2.07 (m, 1H), 2.01 (t, *J* = 3.2 Hz, 1H), 1.99 (m, 2H), 1.97 – 1.85 (m, 2H), 1.82 (dd, *J* = 12.5, 3.5 Hz, 1H), 1.50 (m, 1H), 1.49 (s, 3H), 1.25 (s, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 184.5, 152.8, 151.7, 135.3, 131.2, 127.0, 123.2, 121.5, 113.1, 86.3, 61.5, 58.6, 54.8, 50.7, 49.1, 39.4, 38.2, 30.2, 27.6, 27.3, 22.1, 21.5. IR (film): cm^{-1} 3406, 3249, 2973,

2933, 2876, 2819, 1713, 1578, 1362, 1083, 800. HRMS (ESI, m/z) calcd for $C_{22}H_{25}N_4O_2$ $[M+H]^+$ 377.1978, found 377.1993.

^{1}H NMR (600 MHz, DMSO-d_6) δ 7.59 (s, 1H), 7.55 (m, 2H), 7.47 (t, J = 7.6 Hz, 1H), 7.32 (t, J = 7.4 Hz, 1H), 4.45 (s, 1H), 2.99 (td, J = 8.6, 2.6 Hz, 1H), 2.78 (d, J = 13.5 Hz, 1H), 2.28 (m, 2H), 1.91 (m, 2H), 1.83 (m, 1H), 1.78 (d, J = 13.5 Hz, 1H), 1.75 (m, 1H), 1.64 (dd, J = 12.5, 3.0 Hz, 1H), 1.61 – 1.47 (m, 2H), 1.37 (s, 3H). 1.12 (s, 3H). ^{13}C NMR (150 MHz, DMSO- d_6) δ 185.5, 152.6, 150.9, 135.9, 130.7, 126.6, 123.2, 120.7, 113.0, 85.3, 60.2, 58.0, 54.0, 49.8, 49.0, 36.3, 30.4, 29.5, 26.9, 26.4, 21.6, 21.1.

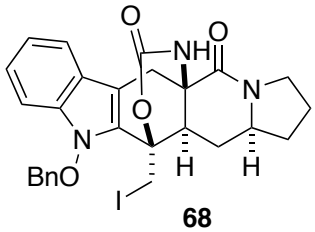


67

To a solution of olefin **35** (100 mg, 0.18 mmol) in acetone (2.65 mL) and H_2O (0.26 mL) was added *N*-methylmorpholine *N*-oxide monohydrate (49 mg, 0.36 mmol) followed by OsO_4 (2% solution in H_2O, 100 μL, 0.008 mmol). The reaction was stirred overnight at room temperature, and was quenched with saturated aqueous $Na_2S_2O_3$ (5 mL). The mixture was extracted with EtOAc (3 x 10 mL), dried over

$MgSO_4$, and concentrated to afford the crude diol. To a solution of ¾ of this crude material in CH_2Cl_2 (4 mL) was added TFA (0.07 mL, 0.91 mmol). The reaction was stirred for 1 hour, and was quenched upon the addition of saturated aq. $NaHCO_3$ (5 mL). The mixture was extracted with CH_2Cl_2 (3 x 10 mL), the combined organic layers were dried ($MgSO_4$) and concentrated. Purification by preparative TLC (1:1 EtOAc/CH_2Cl_2) afforded cyclic carbamate **67** as a white solid (29 mg, approximately 45% yield overall).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.51-7.54 (m, 3 H), 7.42 (m, 3 H), 7.38 (d, *J* = 8.2 Hz, 1 H), 7.31 (t, *J* = 7.5 Hz, 1 H), 7.17 (t, *J* = 7.5 Hz, 1 H), 5.57 (br s, 1 H), 5.18 (m, 2 H), 4.92 (d, *J* = 11.9 Hz, 1 H), 4.09 (d, *J* = 11.9 Hz, 1 H), 3.63 (m, 1 H), 3.52 (m, 2 H), 2.87 (d, *J* = 16.3 Hz, 1 H), 2.75 (m, 1 H), 2.49 (m, 1 H), 2.21 (m, 1 H), 2.08 (m, 1 H), 1.85 (m, 1 H), 1.58 (m, 1 H), 1.45 (q, *J* = 12.2 Hz, 1 H). ^{13}C NMR (150 MHz, $CDCl_3$): δ 167.6, 151.8, 135.1, 133.9, 130.1, 129.7, 129.1, 128.4, 124.8, 122.7, 121.3, 119.6, 110.4, 109.7, 80.8, 77.8, 60.0, 58.5, 57.1, 45.8, 38.8, 33.8, 33.6, 25.1, 22.6. LRMS: ESI(+)-MS for $C_{27}H_{27}N_3O_5$ *m/z*: 496.2 $[M+Na^+]^+$.



68

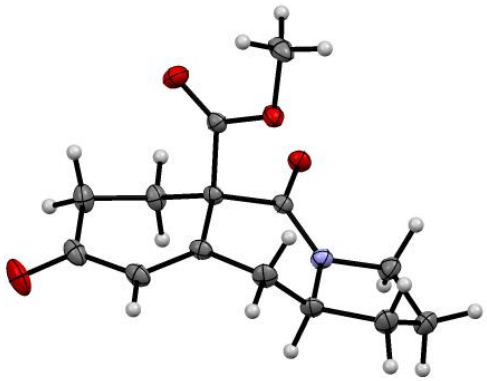
To a solution of **35** (20.0 mg, 0.036 mmol) in CH_2Cl_2 (1 mL) was added I_2 (20.0 mg, 0.079 mmol). The reaction was stirred for 15 minutes at room temperature, and was quenched upon the addition of a solution of saturated aq. $NaHCO_3$ and $Na_2S_2O_3$ (3 mL, 1:1). The mixture was diluted with CH_2Cl_2, and filtered through a phase

separator and concentrated. Purification by preparative TLC (1:1 CH_2Cl_2/Et_2O) afforded **68** as a white solid (10.4 mg, 50% yield).

^{1}H NMR (600 MHz, CD_2Cl_2) δ 7.55 (m, 3 H), 7.44 (m, 4 H), 7.31 (t, *J* = 7.8 Hz, 1 H), 7.16 (t, *J* = 7.8 Hz, 1 H), 5.51 (s, 1 H), 5.19 (d, *J* = 9.5 Hz, 1 H), 5.10 (d, *J* = 9.5 Hz, 1 H), 4.74 (d, *J* = 10.5 Hz, 1 H), 3.59-3.66 (m, 3 H), 3.48-3.52 (m, 2 H), 2.9 (m, 1 H), 2.94 (d, *J* = 16.3 Hz, 1 H), 2.28 (m, 2 H), 2.07 (m, 1 H), 1.87 (m, 1 H), 1.59 (m, 1 H), 1.40 (m, 1 H). HSQC ^{13}C NMR correlations (150 MHz, CD_2Cl_2) 129.9, 129.0, 124.3, 120.7, 119.5, 80.0, 57.5, 53.3, 45.1, 40.7, 33.6, 33.5, 24.0, 22.3, 2.8. LRMS: ESI(+)-MS for $C_{27}H_{26}IN_3O_4$ *m/z*: 606.1 $[M+Na^+]^+$.

Crystal data and structure refinement for compound 19

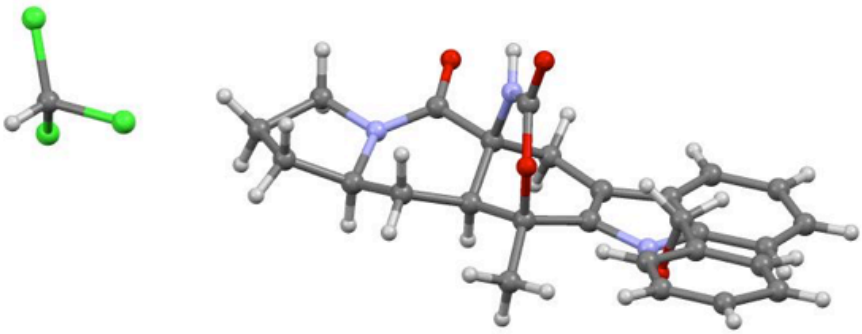
Empirical formula: $C_{14}H_{17}NO_4$
Formula Weight: 263.28
Temperature: 130(2) K
Wavelength: 0.71073 Å
Crystal system, space group: Triclinic, P -1
Unit Cell Dimensions: a = 6.9183(8) Å alpha = 68.2765 (13) deg.
b = 9.7324 (11) Åbeta = 79.1821 (15) deg.
c = 10.5986 (12) Å gamma = 79.3466 (15) deg.
Volume: 645.93 (13) A^3
Z, Calculated Density: 2, 1.354 Mg/m^3
Absorption coefficient: 0.99 mm^-1
F(000): 280
Crystal size: 0.77 x 0.56 x 0.44 mm
Theta range for data collection: 2.086 to 30.489 deg.
Limiting indices: -9<=h<=9, -13<=k<=13, -15<=l<=15
Relections collected / unique: 10400 / 3912 [R(int) = 0.0282]
Completeness to theta = 25.242 99.9%
Absorption correction: Empirical
Max. and min. transmission: 0.7461 and 0.6568
Refinement method: Full-matrix least-squares on F^2
Data / restraints / parameters: 3912 / 0 / 173
Goodness-of-fit on F^2: 1.017
Final R indices [I>2sigma(I)]: R1 = 0.375, wR2 = 0.1070
R indices (all data): R1 = 0.0398, wR2 = 0.1104
Extinction coefficient: n/a
Larges diff. peak and hole: 0.490 and -0.274 e.A^-3



19

Crystal data and structure refinement for compound 40

Empirical formula: $C_{28}H_{28}Cl_3N_3O_4$
Formula Weight: 576.88
Temperature: 130(2) K
Wavelength: 0.71073 Å
Crystal system, space group: Triclinic, P -1
Unit Cell Dimensions: a = 11.434 (2) Å alpha = 70.709 (3) deg.
b = 11.498 (3) Å beta = 64.880 (3) deg.
c = 12.393 (3) Å gamma = 66.319 (3) deg.
Volume: 1325.6 (5) Å ^3
Z, Calculated Density: 2, 1.445 Mg/m^3
Absorption coefficient: 0.387 mm^-1
F(000): 600
Crystal size: 0.37 x 0.17 x 0.15 mm
Theta range for data collection: 1.850 to 30.724 deg.
Limiting indices: -16<=h<=16, -16<=k<=16, -17<=l<=17
Relections collected / unique: 19653 / 8081 [R(int) = 0.0353]
Completeness to theta = 25.242 98.8%
Absorption correction: Empirical
Max. and min. transmission: 0.7461 and 0.6727
Refinement method: Full-matrix least-squares on F^2
Data / restraints / parameters: 8081 / 0 / 344
Goodness-of-fit on F^2: 1.111
Final R indices [I>2sigma(I)]: R1 = 0.1746, wR2 = 0.4746
R indices (all data): R1 = 0.1935, wR2 = 0.4843
Extinction coefficient: n/a
Larges diff. peak and hole: 2.883 and -1.865 e. Å ^-3



40

REFERENCES

1. Park, S. H.; Kang, H. J.; Ko, S.; Park, S.; Chang, S. *Tetrahedron: Asymmetry* **2001**, *12*, 2621-2624.
2. Johnson, C. R.; Adams, J. P.; Braun, M. P.; Senanayake, C. B. W.; Wovkulich, P. M.; Uskokovic, M. R. *Tetrahedron Lett.* **1992**, *33*, 917-918.
3. Brewster, A. G.; Broady, S.; Glenn, E.; Hermitage, S. A.; Hughes, M.; Moloney, M. G.; Woods, G. *Lett. Org. Chem.* **2005**, *2*, 21-24.
4. Shiori, T.; Nonomiya, K.; Yamada, S. J. *J. Am. Chem. Soc.* **1972**, *94*, 6203-6205.
5. Peng, F.; Dai, M.; Angeles, A. R.; Danishefsky, S. J. *Chem. Sci.* **2008**, *130*, 13765-13770.
6. Fischer, E.; Hess, O. *Berichte der Deutschen Chemischen Gesellschaft* **1883**, *17*, 559-568.
7. Kozmin, S. A.; Rawal, V. H. *J. Am. Chem. Soc.* **1998**, *120*, 13523-13524.
8. Nicolaou, K. C.; Estrada, A. A.; Lee, S. H.; Freestone, G. C. *Angew. Chem. Int. Ed.* **2006**, 5364-5368.
9. Belley, M.; Beaudoin, D.; St.-Pierre, G. *Synlett* **2007**, *19*, 2999-3002.
10. Somei, M.; Noguchi, K.; Yamada, F. *Heterocycles* **2001**, *55*, 1237-1240.

11. Yoshida, K.; Goto, J.; Ban, Y. *Chem. Pharm. Bull.* **1987**, *35*, 4700-4704.

12. Zaimoku, H.; Hatta, T.; Taniguchi, T.; Ishibashi, H. *Org. Lett.* **2012**, *14*, 6088-6091.

13. Sakai, S.-I.; Kubo, A.; Katsuura, K.; Mochinaga, K.; Ezaki, M. *Chem. Pharm. Bull.* **1972**, *20*, 76-81.

14. Kametani, T.; Takahashi, K.; Kigawa, Y.; Ihara, M.; Fukumoto, K. *J. Chem. Soc., Perkin Trans. 1* **1977**, 28-31.

15. Tebbe, F. N.; Parhall, G. W.; Reddy, G. S. *J. Am. Chem. Soc.* **1978**, *100*, 3611-3613.

16. Reetz, M. T.; Westerman, J.; Kyung, S.-H. *Chem. Ber.* **1985**, *1050-1057.*

17. Oppolzer, W.; Godel, T. *J. Am. Chem. Soc.* **1978**, *100*, 2583-2584.

18. For a related transformation previously observed in natural products synthesis: Jiricek, J.; Blechert, S. *J. Am. Chem. Soc.* **2004**, *126*, 3534-3538.

19. Knölker, H.-J.; Braxmeier, T.; Schlechtingen, G. *Angew. Chem. Int. Ed. Eng.* **1995**, *34*, 2497-2500.

20. Kilényi, S. N. "Sodium 4,6-Diphenyl-1-oxido-2-pyridone." *e-Eros Encyclopedia of Reagents for Organic Synthesis* **2001**.

21. Katritzky, A. R.; Cook, M. J.; Brown, S. B.; Cruz, R.; Millet, G. H.; Anani, A. *J. Chem. Soc., Perkin Trans. 1.* **1979,** 2493-2499.

22. Miyoshi, T.; Matsuya, S.; Tsugawa, M.; Sato, S.; Ueda, M.; Miyata, O. *Org. Lett.* **2013**, *15*, 3374-3377.

23. Chen, D. X.; Ho, C. M.; Wu, Q. Y. R.; Wu, P. R.; Wong, F. M.; Wu, W. *Tetrahedron Lett.* **2008**, *49*, 4147-4148.

24. Kornblum, N.; Jones, W. J.; Anderson, G. J. *J. Am. Chem. Soc.* **1959**, *81*, 4113-4114.

25. Braun, N. A.; Bray, J. D.; Ciufolini, M. A. *Tetrahedron Lett.* **1999**, *40*, 4985-4988

26. Colomer, I.; Batchelor-McAuley, C.; Odell, B.; Donohoe, T. J.; Compton, R. G. *J. Am. Chem. Soc.* **2016**, *138*, 8855-8861.

27. Hecht, S. M.; Kozarich, J. W. *J. Chem. Soc., Chem. Commun.* **1973**, 387-388.

28. Durand, J.-O.; Larchevêque, M.; Petit, Y. *Tetrahedron Lett.* **1998**, *39*, 5743-5746.

29. Ji, N.-Y.; Liu, X.-H.; Miao, F.-P.; Qiao, M.-F. *Org. Lett.* **2013**, *15*, 2327-2329.

30. Levinson, A. M. *Org. Lett*. **2014**, *16*, 4904-4907.

31. Kushida, N.; Watanabe, N.; Okuda, T.; Yokoyama, F.; Gyobu, Y.; Yaguchi, T. *J. Antibiot.* **2007**, *60*, 667-673.

32. Berlin, M.; Boyce, C. W.; de Lera Ruiz, M. *J. Med. Chem.* **2011**, *54*, 26-53.

33. Bian, Z.; Marvin, C. C.; Pettersson, M.; Martin, S. F. *J. Am. Chem. Soc.* **2014**, *136*, 14184-14192.

34. Iwama, T.; Birman, V. B.; Kozmin, S. A.; Rawal, V. H. *Org. Lett.* **1999**, *1*, 673-676.

Chapter 3: Introduction to the Total Chemical Synthesis and Folding of All-L and All-D KRas(G12V)

Targeting Ras Signaling Through the Means of Chemical Protein Synthesis

The RAS oncogene encodes small GTPases involved in cellular signaling, controlling a variety of phosphorylation cascades involved in cellular growth and target gene transcription. Scolnick and co-workers initially identified Ras in 1979 as a 21 kDa protein encoded by the Harvey and Kirsten rat sarcoma viruses in transformed cells.[1,2] In 1982, three independent laboratories identified sequence homologs of these cancer-inducing genes within human bladder and lung cancer cells, thereby establishing RAS as an endogenous human oncogene.[3,4,5] Since these seminal discoveries, the four human isoforms of this oncogenic protein (HRas, NRas, KRas-4A, and KRas-4B) have been implicated in over 30% of all human cancers, with higher prevalence in certain cancers of particular mortality, including 71% of pancreatic cancer cases (**Table 1**).[6] Of the four human Ras isoforms, KRas-4B is responsible for a vast majority of these cancers.

Table 1: Selected frequency of mutated Ras isoforms in cancer (adapted from Ref. 6 based on the Catalogue of Somatic Mutations in Cancer (COSMIC) version 67)

Selected Frequency of Mutated Ras Isoforms in Cancers

Tissue	HRas (%)	KRas (%)	NRas (%)	total Ras%
Pancreas	0%	71%	<1%	71%
Lung	<1%	19%	<1%	20%
Colon	1%	35%	6%	42%
Small Intestine	0%	35%	<1%	35%
Biliary Tract	0%	26%	2%	28%
Skin	1%	1%	18%	20%
Cervix	9%	8%	2%	19%
Urinary Tract	10%	5%	1%	16%

The functional significance of Ras in cellular signaling was established within the first years following the initial discoveries of this human gene homolog. Ras is a GTP-

binding protein, which controls intracellular signaling cascades, as governed by activation of receptor tyrosine kinases (RTKs), such as extracellular epidermal growth factor (EGF) binding.[7] This extracellular signaling recruits proteins known as guanine nucleotide exchange factors (GEFs), such as the Son of Sevenless protein (SOS), which facilitate nucleotide exchange of GDP-bound Ras to allow for GTP binding (**Figure 1**).[8,9] Conversely, although Ras itself is a GTPase, its innate hydrolysis of GTP is slow, and requires the aid of GTPase activating proteins (GAPs) to facilitate hydrolysis of GTP to GDP. [10,11] The structural differences between GDP- and GTP-bound states of Ras govern its effector binding and signal transduction. This renders Ras as a crucial molecular switch at the intersection between the extracellular environment and cytosolic phosphorylation cascades that regulate transcription and growth. In its GTP-bound state, Ras binds several effector proteins, initiating biological outcomes including cellular proliferation, survival, differentiation and apoptosis. Most significantly as tied to its carcinogenesis, Ras mediates signal transduction primarily via the BRaf-MEK-ERK and PI3K-AKT-mTOR pathways, each independently implicated in cancer.[12,13]

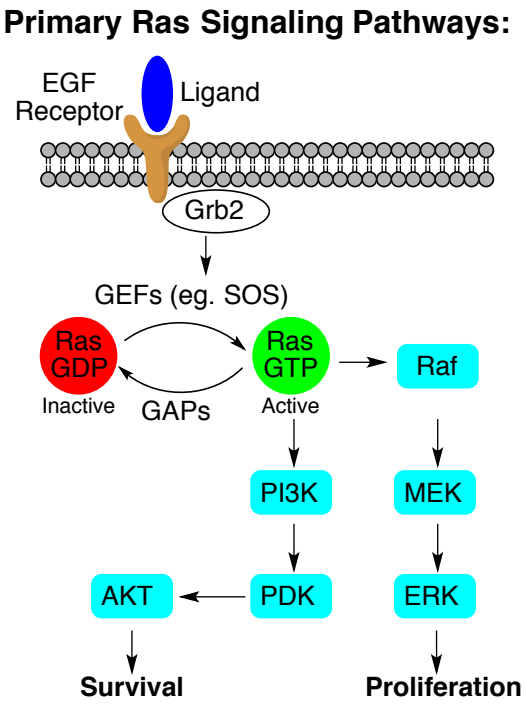


Figure 1: Primary Ras signaling pathways involved in cancer.

Structurally, Ras proteins are comprised of two main peptide regions – the catalytic G-region (residues 1-166) and the hypervariable region (residues 167-188 or 189) (**Figure 2**). This C-terminal hypervariable region is a non-conserved polypeptide portion linking the catalytic region of the protein to a lipidated tail for lipid-bilayer anchoring. This region terminates in a –CAAX motif (a four-residue sequence comprised of: Cys-Aliphatic-Aliphatic-Any Amino Acid) for recognition and post-translational modification. Variation within this region is responsible for the differential localization and processing of Ras isoforms within the cell.[14,15] Within the G-region of Ras, residues 1-86 are entirely conserved across all Ras isoforms.[16] The GTP-binding active site of Ras (and contact points for effector binding) consists of three main regions: 1) the P-loop, which binds to the β-phosphate of GTP, 2) the switch I domain, which binds the γ-phosphate of GTP, 3) the switch II domain, which also binds the γ-phosphate of GTP, is

important for nucleoside recognition, and contains Gln61 as a critical catalytic residue for GTP hydrolysis.[17,18] Conformational changes in the Switch I and II regions largely control the "on" and "off" states of Ras signaling. Although several sequence mutations are tolerated with regard to Ras function, specific single point mutations are remarkably capable of transforming proto-oncogenic Ras into a constitutively active oncogenic form by obstructing effector binding and by reducing intrinsic GTPase activity.[19-22] Mutations at Gly12 or Gly13 within the P-loop sterically impede GAP interactions for GTP hydrolysis, therefore locking Ras in a GTP-bound "on" state.[23] Additionally, Gln61 in the switch II region is important for coordination of a water molecule in GTP hydrolysis, and mutation of this residue can also lead to cancer development.[24]

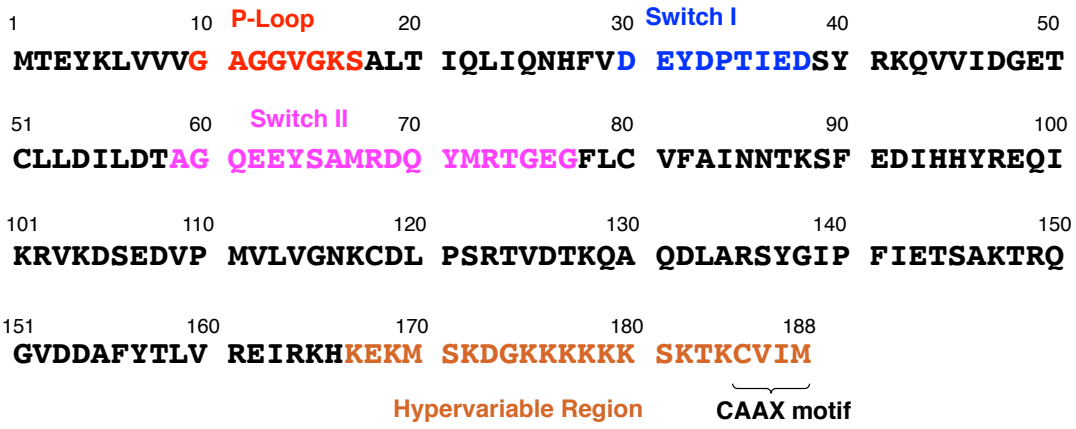


Figure 2: Primary structure of KRas-4B with highlighted key structural regions

Strategies for Targeting Ras

The direct inhibition of Ras is a logical approach to targeting Ras-driven cancers. However, most approaches report only limited success. **Figure 3** highlights general strategies for targeting Ras signaling.[25,26] These strategies include: 1) direct inhibition of RasGTP binding to its effectors, 2) prevention of RasGTP formation by inhibiting Ras-

GEF interactions or by stabilizing RasGDP. For clarity, this thesis only highlights approaches for the direct targeting of Ras. However, significant research centered on *indirect* targeting of Ras by inhibiting upstream or downstream kinase pathways is equally important but will not be discussed here in detail. These strategies are complicated, as inhibition of downstream targets results in a variety of negative feedback loops, which can cause paradoxical activation of Ras pathways, and inhibition of upstream targets such as the EGFR often lead to resistance.[27,28] Additionally, particular effort has focused on the inhibition of Ras post-translational modification pathways to prevent anchoring of the protein at the cellular lipid bilayer, largely through the development of farnesyltransferase inhibitors (FTIs) and geranylgeranyltransferase 1 inhibitors (GGTIs) among others.[29] Unfortunately, despite promising preclinical data surrounding FTIs, this strategy has generally proven futile in clinical trials with regard to KRas-driven cancers.[30] Other promising strategies in cancer treatment, such as synthetic lethal screening, may hold promise for treating Ras-driven cancers, but successes in this area have not been realized.[31]

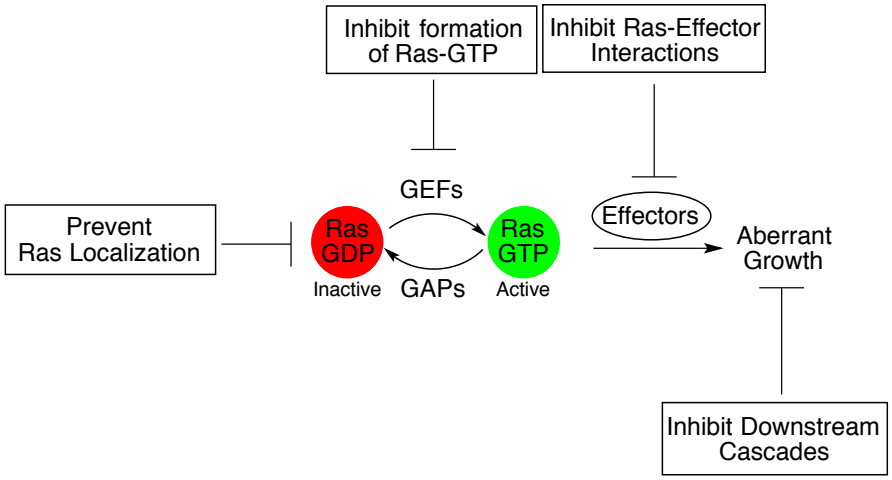


Figure 3: Therapeutic strategies for targeting Ras-driven cancers

Direct targeting of Ras using small molecules is particularly challenging. The primary reason for this is the lack of known or available hydrophobic binding pockets within the protein structure, with the exception being the GTP-binding active site. Unfortunately, targeting of the active site of Ras is largely futile, due to the unyielding affinity of Ras for its nucleotide substrate (picomolar range) and the high concentration of GTP within cells. Ras binding with GEFs or effectors are also governed by large protein-protein interactions (PPIs), targeting of which via a small molecule approach is difficult.[32,33] In recent years, advances in methods such as fragment-based drug screening and *in silico* methods have enabled identification of shallow binding pockets and small molecules that disrupt these PPIs.[25] Despite these developments, the majority of identified small molecules lack the potency required for clinical application.[34] Highlighted in **Figure 4** are representative examples of these identified small molecules and their general mechanisms.[35-38] In a most-promising strategy, work by the Shokat and Gray labs in 2013 identified small molecules that covalently target cysteine in mutant KRas(G12C) (**5** and **6**).[39,40] Although the KRas(G12C) mutant represents a small percentage of all KRas-driven cancers, this mutation is found in a relatively high percentage of lung cancers (7% of lung cancers; 40% of KRas-driven lung cancers), rendering it a therapeutically relevant mutation. In the case of Shokat's inhibitor (**5**), the molecule occupies a previously unidentified allosteric binding pocket, blocking effector binding, and structurally favoring GDP over GTP binding. On the other hand, electrophilic GDP analog **6** covalently binds to KRas(G12C) and occupies the active site, favoring the GDP-bound "off" conformation.

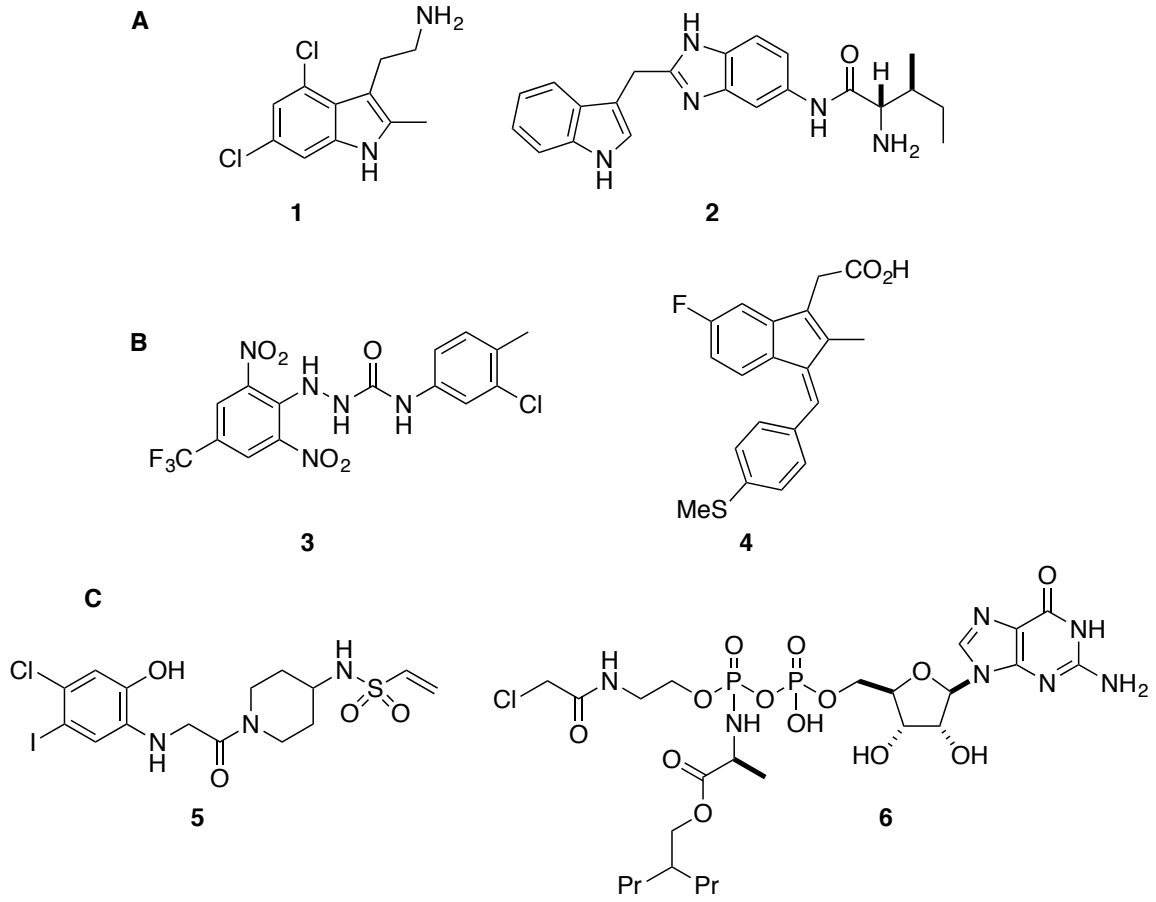


Figure 4: Representative small molecules that target Ras. (**A**) Ras-GEF inhibitors. (**B**) Ras-Raf inhibitors. (**C**) Electrophilic covalent inhibitors of KRas(G12C)

Aside from small molecules, peptides have shown promise in the direct targeting of PPIs, including those associated with Ras biology.[41] Of particular significance, constrained α-helical peptides represent a promising class of cell-penetrating modulators of PPIs with improved bioavailability over unconstrained peptides, and have been recently developed in the context of Ras biochemistry (**Figure 5**). In 2011, a collaborative venture between the Arora and Bar-Sagi labs identified a cell-penetrant 16-mer α-helical peptide based on the structure of the SOS αH helix (**7**). This constrained peptide, designed using hydrogen bond surrogate (HBS) tethering, directly inhibits the SOS-Ras interaction.[42,43] It was demonstrated that inhibition by this peptide decreases

levels of RasGTP upon EGFR activation, and attenuates Ras signaling as marked by decreased phosphorylation levels of downstream ERK. In 2015, Walensky and co-workers at the Dana-Farber Cancer Institute identified a cell-penetrating hydrocarbon stapled peptide based on a similar sequence, but with improved binding affinity (**8**).[44,45] Moreover, this peptide was shown to impair the viability of several cancer cell lines containing oncogenic mutations at Gly12, Gly13, and Gln61.

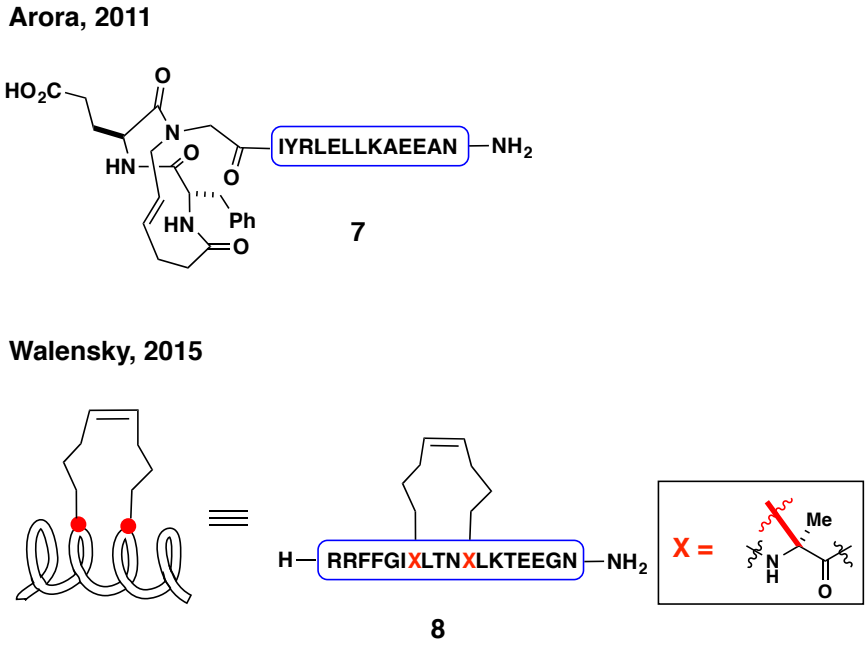


Figure 5: Synthetic α-helical peptide inhibitors of Ras-SOS interactions.

Our Approach Toward Targeting Ras: Mirror Image Yeast Surface Display (YSD) to Identify all-D Amino Acid Peptide Inhibitors of Oncogenic Ras

Recently, Verdine and co-workers developed a class of miniproteins that bind to the Ras effector domain with high affinity and inhibit Ras-effector interactions.[46] These peptides were discovered by screening unbiased combinatorial libraries with yeast surface display (YSD), demonstrating that it is possible to identify high-affinity Ras

ligands from naïve libraries.[47] With these results in mind, we wondered if mirror-image YSD could be used to identify high-affinity Ras binders composed of all D-amino acids.[48] Unnatural, D-residue peptides have been shown to garner augmented bioavailability over their L-residue peptide counterparts and may be less prone to eliciting an immune response.[49] Mirror-image display has successfully afforded useful peptide ligands in the context of various disease types, including HIV/AIDS, Alzheimers' Disease, and cancer. **Figure 6** highlights the basic principles of mirror image display. Crucial to this method is the chemical synthesis of an all-D amino acid form of the protein target of interest. The power of combinatorial biological expression libraries, capable of expressing up to 10^{13} different polypeptides, can then be exploited by screening against the all-D protein. Upon selection and structure determination of ligands, the all-D amino acid form of the ligand can be prepared to bind the biological target interest.

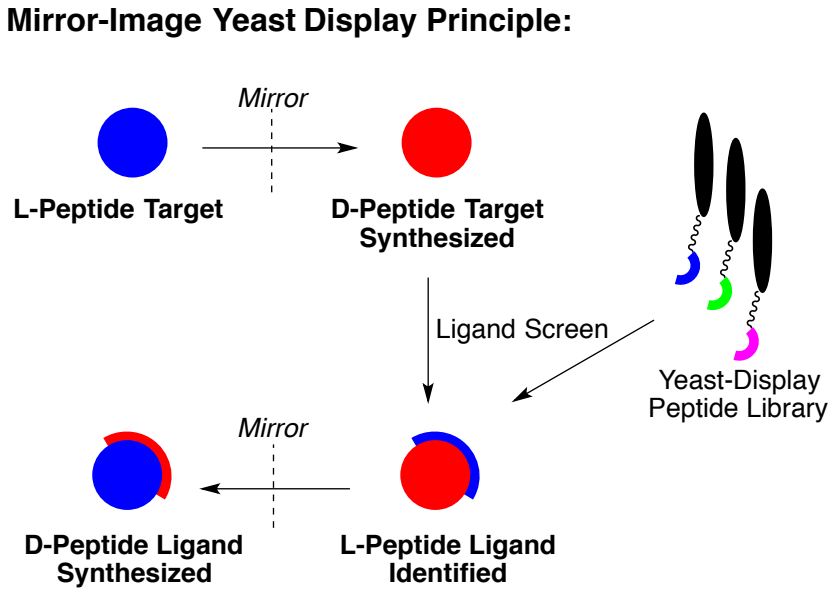


Figure 6: Principle concepts behind mirror-image yeast surface display (YSD)

In a related example, Lu and co-workers utilized mirror-image phage display to identify all-D dodecapeptide inhibitors of the p53-MDM2 interaction.[50,51] Notably, MDM2 is an intracellular target and serves as a negative regulator of the tumor-suppressor protein, p53. Overexpression of MDM2 is implied in cancers, and p53 inhibition by MDM2 disrupts its regulatory biological functions in the cell cycle and prevention of tumor growth. For use in mirror image phage display, all-D MDM2[25-109] was prepared using a combination of solid-phase peptide synthesis (SPPS) and Native Chemical Ligation (NCL) strategies. The researchers identified several α-helical all-D amino acid sequences that bind to the p53-binding domain of MDM2. The most potent sequence identified from their libraries, ^{D}PMI-γ (sequence: H-DWWPLAFEALLR-NH_2, consisting of all-D residues), competes with p53[15-29] binding to MDM2 with a K_d value of 53 ± 6 nM. Subsequent structural optimization of sequences found through this method yielded even more potent binders with sub-nanomolar affinity.[52]

The intracellular nature of regulatory protein inhibition presents a separate challenge for peptide-based therapies, which are frequently unable to traverse the cellular membrane. Despite potent binding affinity, ^{D}PMI-γ is incapable of passing through the lipid bilayer. However, the authors were successful in delivering the peptide into cells and observing cytotoxicity by encapsulation of the D-peptide within liposomes. *In vivo* experiments resulted in inhibition of glioblastoma tumor growth in mouse models. Additionally, Pentelute and co-workers demonstrated that these D-peptides can also be delivered intracellularly by utilizing proteins derived from *Bacillus anthracis* involved in transport of Anthrax toxins into cells. Conjugation of an MDM2-p53 inhibitor to the

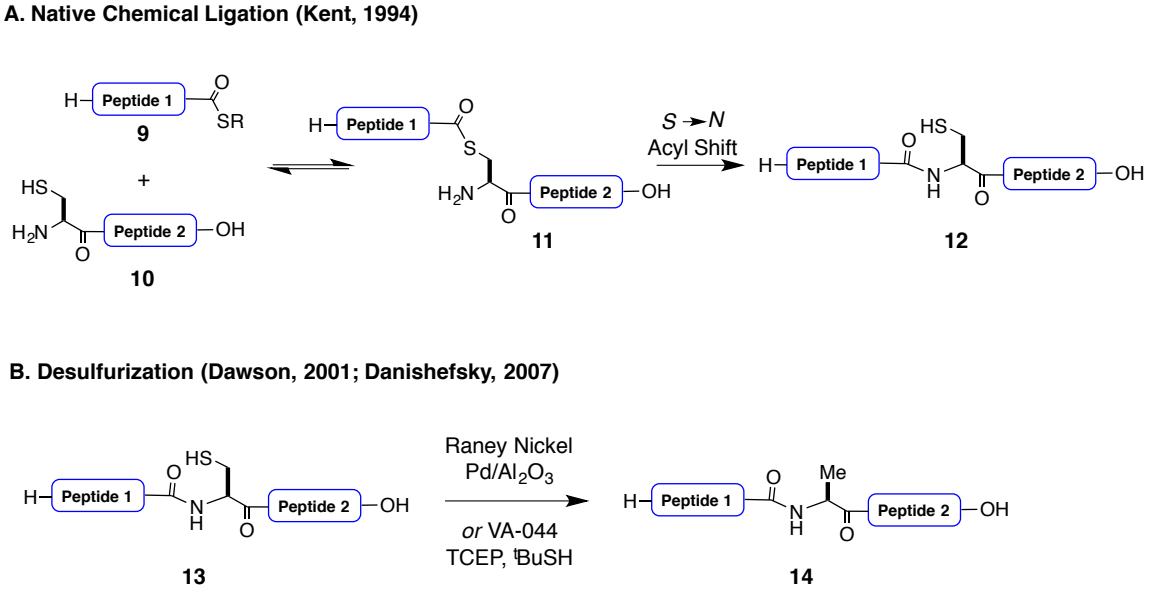
lethal factor (LF) of anthrax toxin enables transport across the cell membrane.[53] Although intracellular delivery of peptide therapeutics remains a global challenge, these results highlight that D-peptides have potential as therapeutics, and novel technologies for their delivery are becoming available.

Methods for the Synthesis of Large Polypeptides and Proteins

Critical to our strategy for targeting KRas is the total chemical protein synthesis and proper folding of an all-D residue protein variant. Since a seminal report by R. Bruce Merrifield in 1963, solid-phase peptide synthesis (SPPS) has become state of the art for peptide and protein synthesis – a development for which Merrifield was awarded the Nobel Prize in 1984.[54] The development of SPPS was groundbreaking, enabling the use of excess reagents and a single reaction flask to achieve high yielding syntheses of polypeptides and proteins previously unattainable by classical solution-based methods.[55] Modern SPPS is a general and efficient means to generate peptide fragments up to approximately 40-60 amino acids, with sequence-specific determination of the overall efficiency.[56] The advent of powerful methods for the synthesis of longer peptide sequences via ligation strategies further enables our ability to address biological questions. Most notably, the development of Native Chemical Ligation (NCL) by Kent and co-workers in 1994 has enabled the synthesis of functional proteins well beyond lengths attainable by SPPS.[57] **Scheme 1A** highlights the principle of the NCL reaction between two appropriate peptide partners. The reaction proceeds through an initial reversible transthioesterification reaction between a C-terminal thioester (**9**) and an N-

terminal cysteine residue (**10**). The intermediary ligated thioester (**11**) subsequently undergoes S → N acyl transfer to irreversibly form the native peptide bond (**12**).

Due to the relative scarcity of cysteine within proteins, our laboratory and others have worked to extend the utility of NCL. In the first permutations of alternative solutions, a non-natural Cys residue can be introduced, and subsequently desulfurized post-ligation to provide a native Ala residue (**Scheme 1B**). Dawson and co-workers demonstrated that following NCL at a non-native cysteine residue as the nucleophilic peptide reactant, the β-thiol can be desulfurized through the action of Pd/Al_2O_3 or Raney nickel, yielding an alanine residue.[58,59] In 2007, our laboratory demonstrated that desulfurization can be accomplished under free radical conditions, utilizing VA-044 or V-50 as radical initiators at ambient or slightly elevated temperatures in the presence of tris(2-carboxyethyl)phosphine (TCEP) and a thiol co-catalyst (2-methyl-2-propanethiol).[60] Moreover, these conditions innocuous to sensitive residues, such as methionine, which are incompatible with previously reported methods. Our metal-free desulfurization (MFD) method has evolved since these initial reports to include many other amino acids bearing a β-thiol moiety, extending the utility and number of disconnection sites available for NCL.[61]



Scheme 1: (**A**) Native Chemical Ligation (NCL) and (**B**) desulfurization of NCL products

Aside from NCL-based approaches reliant on an initial transthioesterification step, it is instructive to feature other methods for solution-phase peptide ligation that are independent of cysteine. **Figure 7** highlights representative examples of such transformations. Bode and co-workers developed a chemoselective reaction between a C-terminal ketoacid and N-terminal hydroxylamine or hydroxylamine derivative.[62] This ketoacid-hydroxylamine transformation (termed *KAHA* ligation) takes place through the intermediacy of an oxime, followed by loss of CO_2 and H_2O to yield the desired amide bond. Highlighted is the chemoselective ligation between **15** and **16** in the synthesis of GLP-1[7-36] **17**, along with a small amount of epimer at Glu21.[63] Another well-applied method for the synthesis of long polypeptide chains is the "thioester method," entailing the Ag(I)-mediated activation of *C*-terminal alkyl thioesters in the presence of any *N*-terminal nucleophile, requiring only partial protection of Lys and Cys residues.[64] One example is showcased in the synthesis of the 90-residue HU-type DNA binding protein

(HBs, **20**). With the *N*-terminus of **18** as well as the 12 additional Lys residues within **18** and **19** orthogonally protected with *tert*-butyloxycarbonyl (Boc) groups, these peptides underwent ligation in the presence of $AgNO_3$ and *N*-hydroxysuccinimide (HOSu) to yield HBs[1-90] **20** in 33% yield after deprotection.[65] In an impressive feat, this method was later applied in the synthesis of a 141-residue glycopeptide containing 6 tandem repeats of a highly glycosylated MUC2 fragment (*not shown*).[66] A complementary approach developed by our laboratory involved the chemoselective amide bond formation between an *ortho*-disulfide phenolic ester **21** and the *N*-terminus of a second peptide fragment **22** under reducing conditions to form glycopeptide **23**.[67] Notably, the alkyl thioester survives this transformation, and the ligated glycopeptide **23** could then be applied in a second Ag(I)-mediated fragment coupling. Sparsely applied for peptides, but mechanistically unique is the Staudinger-type ligation described by Raines and co-workers.[68] Reaction of an (*ortho*-phosphino)arylthioester (**24**) with an azide nucleophile (**25**) generates iminophosphorane intermediate **26**, which undergoes S → N acyl transfer and hydrolysis of the phosphorus-nitrogen bond to generate peptide **27**.

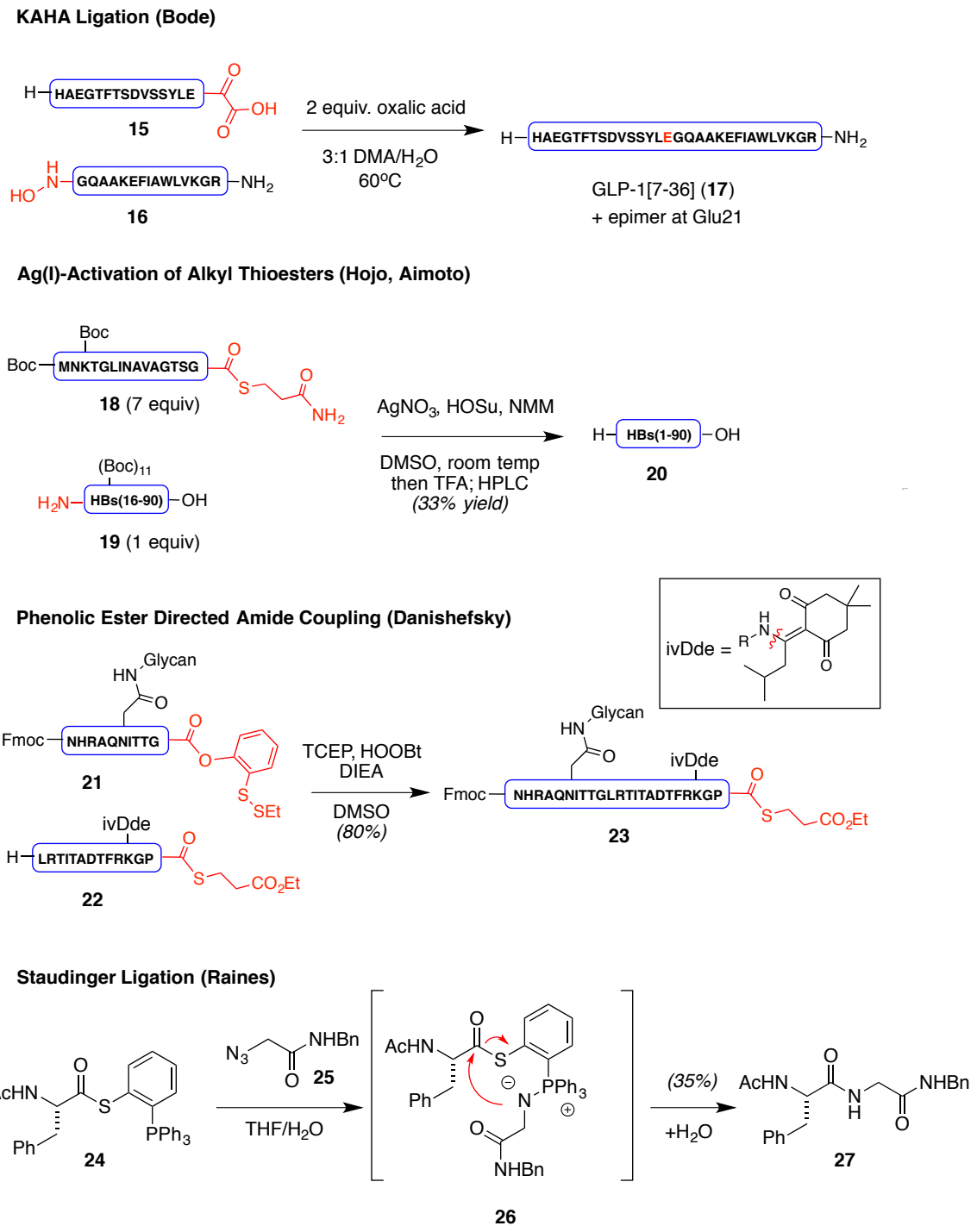


Figure 7: Alternative peptide ligation methods independent of cysteine

In addition to the above described methods for cysteine-free direct amide-bond formation of fully deprotected or minimally protected peptides, a number of strategies to facilitate peptide ligation were developed based on temporal auxiliaries either on the *N*-

terminal or *C*-terminal reacting partner with designed functionality for favorable post-ligation cleavage (**Figure 8**). One such example is the ligation between **28** and **29** through the assistance of a mercaptoarene auxiliary on the nucleophilic fragment.[69] Upon reduction with TCEP, **28** undergoes *ortho*-mercaptoaryl ester rearrangement (OMER), followed by transthioesterification with **29**, and lastly *S* → *N* acyl transfer to yield the desired glycopeptide **30**. This auxiliary was later removed in a 2-step sequence, involving *S*-methylation of the auxiliary followed by deprotection under acidic conditions (*not shown*). To highlight one example of an electrophilic auxiliary, Li and co-workers described an auxiliary based on a salicylaldehyde ester that will undergo chemoselective ligation with *N*-terminal serine or threonine residues.[71] This work is based on previous work by Tam and co-workers, which demonstrated a similar ligation at Ser, Thr and Cys, but without successful auxiliary removal.[72] Shown in **Figure 8** is the application of this method to a synthesis of human erythrocyte acylphosphatase. The reaction between **31** and **32** proceeds via aldehyde capture to form an *N,O*-benzylidine acetal intermediate (*not shown*), which undergoes *O* → *N* acyl transfer to generate pseudoproline **33**. Removal of the auxiliary could then be effected through treatment with TFA to generate the native amide bond (*not shown*).

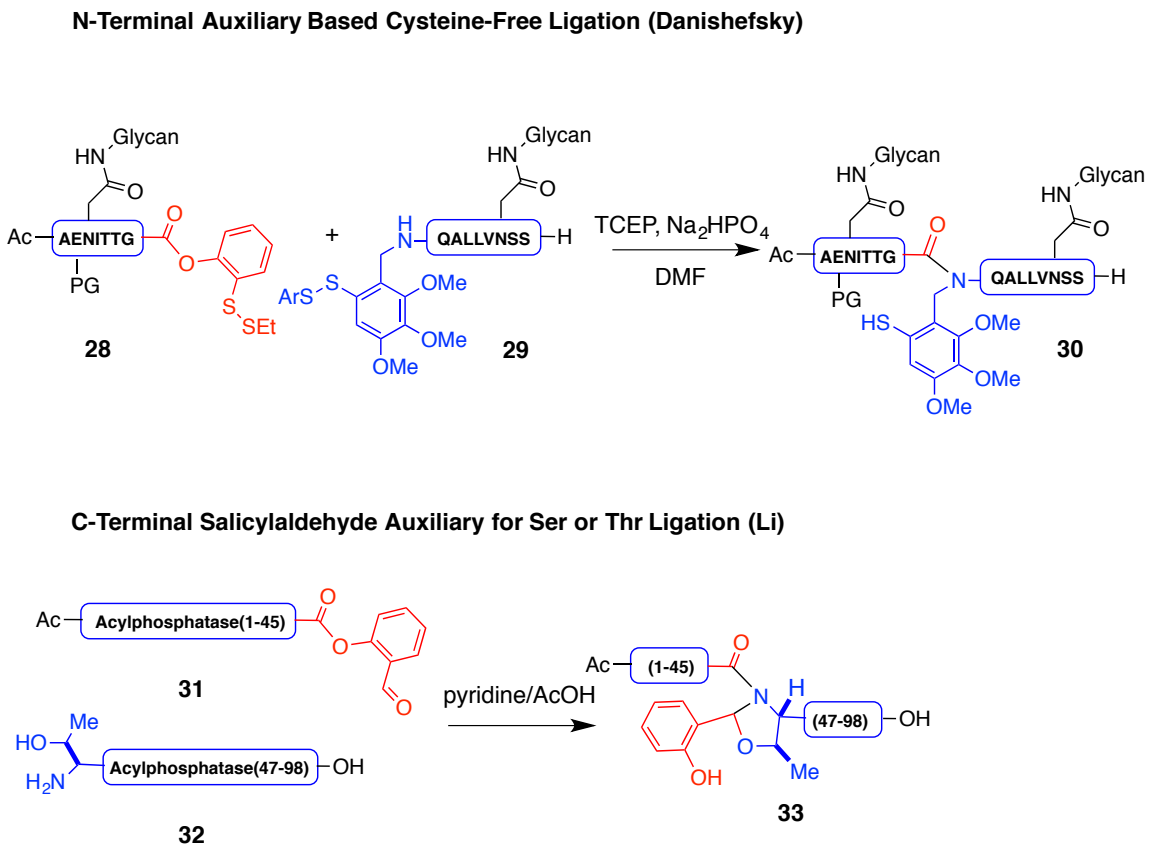


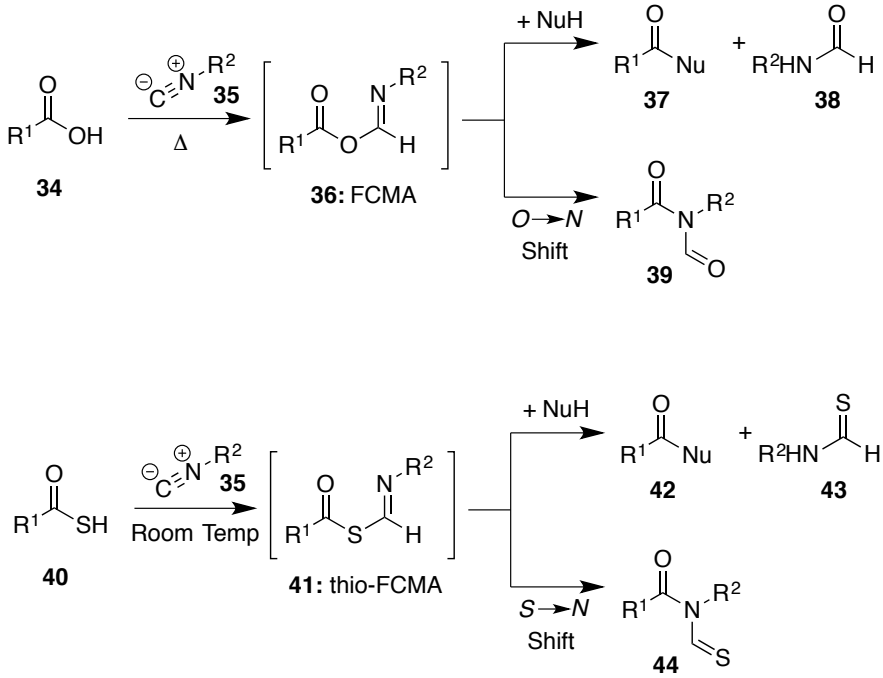
Figure 8: Auxiliary-based cysteine-free ligations

Chemoselective Reactions of Isonitriles with Thiocarboxylic Acids: Application to Peptide and Protein Synthesis

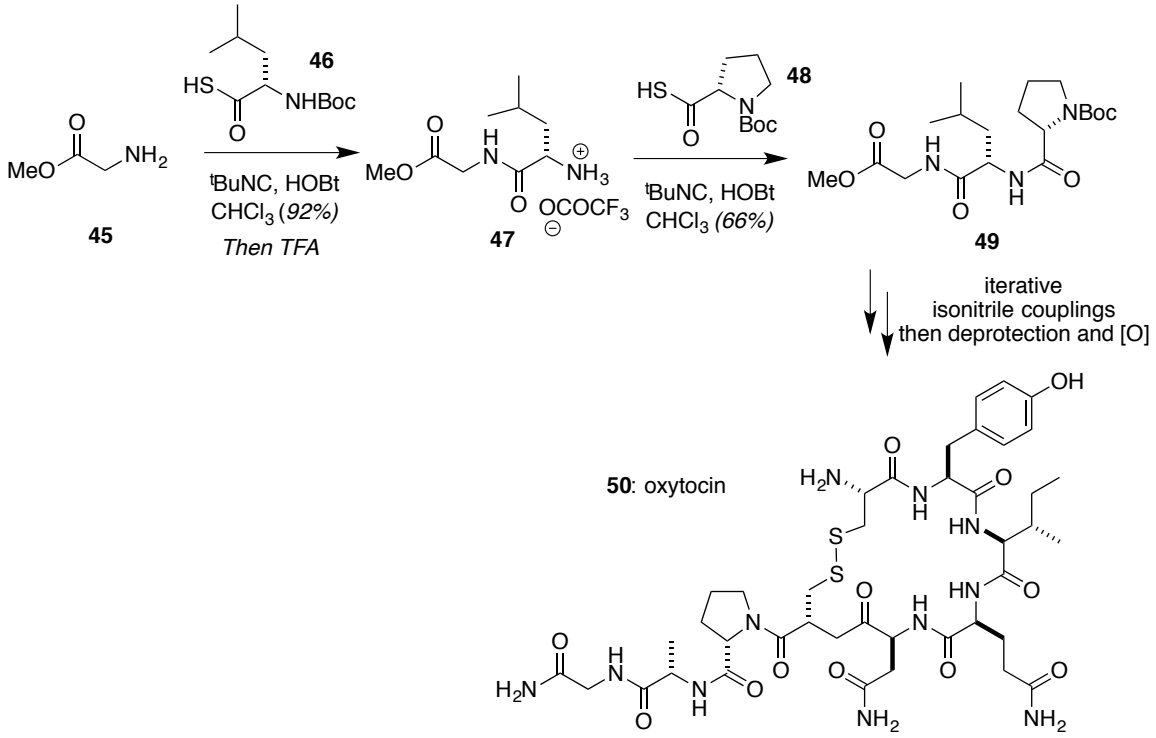
Our laboratory has a long-standing interest in the utility of isonitriles in organic synthesis, particularly as a powerful tool for amide bond synthesis. My work builds upon years of studies that has allowed us to explore isonitrile mediated activation of thioacids for medium- and large peptide ligations. In this thesis, our work in the context of polypeptide synthesis is briefly highlighted. **Scheme 2** shows the reactivity differences of carboxylic acids (**34**) compared to thiocarboxylic acids (**40**) when treated with an isonitrile (**35**). Under thermal conditions, carboxylic acids undergo activation with

isonitriles to generate a presumed formimidate carboxylate mixed anhydride (FCMA, **36**) intermediate. Depending on the steric nature of R^2 on the isonitrile and the presence of an exogenous nucleophile, such as an amine, alcohol, or HOBt, **36** can either undergo acylation with a nucleophile to generate **37**, or alternatively can undergo 1,3-*O* → *N* acyl transfer to produce *N*-formylamide **39**.[72,73] Intriguingly thiocarboxylic acids were found to produce presumed thio-FCMA (**41**) intermediates at ambient temperature, which undergo similar reactivity.[74,75]

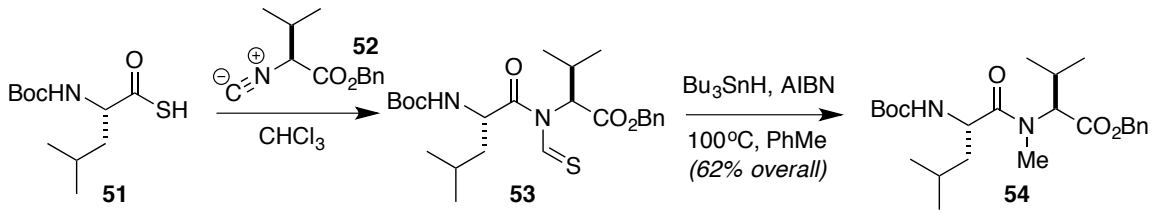
In an application to peptide synthesis, our laboratory demonstrated that amino acid thiocarboxylic acids can undergo efficient amide coupling at room temperature. Adamantyl or *tert*-butyl isocyanides were used as bulky, hindered isonitriles to prevent undesired rearrangements (*i.e.* S → N intramolecular shift, see **Scheme 2**). This strategy was applied in a total synthesis of oxytocin **50** (**Scheme 3**) using iterative isonitrile coupling/deprotection reactions in the C → N direction.[76] It should be mentioned that in addition to amide bond synthesis via isonitrile activation of thiocarboxylic acid reactants, thioacids have previously been utilized as precursors to amides through other reactivity modes with coupling partners such as isocyanates, azides, and sulfonamides.[77,78,79] We have also demonstrated the utility of amino acid isonitriles in the synthesis of *N*-methylated peptides (**Scheme 4**). One example is the coupling of leucine thioacid **51** with valine isocyanide **52** *en route* to cyclosporine A.[80] Reduction of the intermediary *N*-thioformyl dipeptide **53** yielded *N*-methylated dipeptide **54** in 62% overall yield.



Scheme 2: Reactivity of carboxylic acids and thiocarboxylic acids with isonitriles



Scheme 3: Application of iterative isonitrile-activated coupling of thiocarboxylic acids to a total synthesis of oxytocin (**50**)



Scheme 4: Synthesis of *N*-methylated peptides through the means of amino acid isonitriles *en route* to cyclosporine A.

The capacity to chemoselectively activate thiocarboxylic acids under ambient conditions opens the door for the ligation of peptides with minimally protected side-chains. Indeed, we previously demonstrated that thioacids are also activated under oxidative, yet mild, conditions promoted by reactions in DMSO open to air (**Figure 9A)**.[81] Oxidative activation of thioacid **55** in the presence of HOBt yielded ligation product **57** in 82% yield. In an impressive realization of isonitrile-mediated peptide ligation, our laboratory recently demonstrated the coupling of two large peptide fragments via activation of *C*-terminal thioacid **58** and nucleophilic capture with **59** to generate a 101-amino acid sequence **60** after Fmoc removal *en route* toward granulocyte-colony stimulating factor (G-CSF).[81] Although this fragment was obtained in low yield, it is noteworthy that a more conventional approach for assembly via NCL-MFD at an "alanine disconnection" failed to deliver the putative sequence. Currently we are studying the further utility of this method for medium- and large peptide ligations at non-epimerizable residues Gly and Pro. Fortunately, both glycine and proline are relatively abundant amino acids, and such disconnections aid in the complementarity of this approach to NCL.

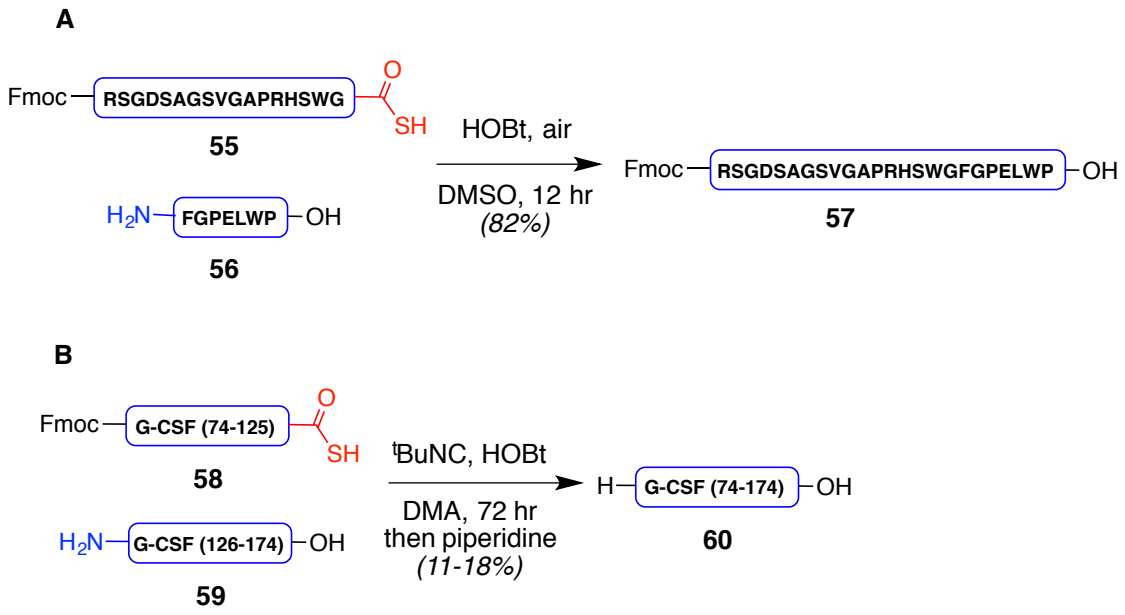


Figure 9: Peptide ligation of C-terminal thioacids (**A**) Oxidative peptide coupling of *C*-terminal thiocarboxylic acids. (**B**) Isonitrile-mediated ligation *en route* toward G-CSF.

Conclusion:

With an understanding of the literature methods for peptide synthesis and previous approaches for targeting Ras, the following section will highlight our synthetic approach to KRas, in which we are targeting a commonly mutated G12V oncogenic isoform. Subsequent biochemical experiments seek to establish the proper folding of Ras in a GTP-bound state, for which GppNHp will be utilized as a non-hydrolyzable GTP analog. Although the synthesis of *wt* HRas was previously described, our synthetic approach towards this target is distinct. Moreover, synthesis of an all-D variant and folding of the mirror-image protein with a mirror-image nucleotide substrate has never been demonstrated. These synthetic studies and subsequent mirror image YSD experiments will seek to identify new, all-D amino acid residue peptides to target Ras-GTP and block effector interactions.

References:

1. Shih, T. Y.; Weeks, M. O.; Young, H. A.; Scolnick, E. M. *Virology* **1979**, *96*, 64-79.
2. Malumbres, M.; Barbacid, M. *Nat. Rev.* **2003**, *3*, 7-13.
3. Der, C. J.; Krontiris, T. G.; Cooper, G. M. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 3637-3640.
4. Parada, L. F.; Tabin, C. J.; Shih, C.; Weinberg, R. A. *Nature* **1982**, *297*, 474-478.
5. Santos, E.; Tronick, S. R.; Aaronson, S. A.; Pulciani, S.; Barbacid, M. *Nature* **1982**, *298*, 343-347.
6. Stephen, A. G.; Esposito, D.; Bagni, R. K.; McCormick, F. *Cancer Cell.* **2014**, *25*, 272-281.
7. Olson, M. F.; Marais, R. *Semin. Immunol.* **2000**, *12*, 63-73.
8. Downward, J.; Riehl, R.; Wu, L.; Weinberg, R. A. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 5998-6002.
9. Bonfini, L.; Karlovich, C. A.; Dasgupta, C.; Banerjee, U. *Science* **1992**, *255*, 603-605.
10. Adari, H.; Lowy, D. R.; Willumsen, B. M.; Der, C. J.; McCormick, F. *Science* **1988**, *240*, 518-521.
11. Bos, J. L.; Rehmann, H.; Wittinghofer, A. *Cell*, **2007**, *129*, 865-877.
12. Khosravi-Far, R.; Der, C. J. *Cancer and Metastasis Rev.* **1994**, *13*, 67-89.
13. Rodriguez-Viciana, P.; Warne, P. H.; Dhand, R.; Vanhaesebroeck, B.; Gout, I.; Fry, M. J.; Waterfield, M. D.; Downward, J. *Nature* **1994**, *370*, 527-532.
14. Wright, L. P.; Philips, M. R. *J. Lipid Res.* **2006**, *47*, 883-891.

15. Hancock, J. F. *Nat. Rev. Mol. Cell. Biol.* **2003**, *4*, 373-384.

16. Bar-Sagi, D. *Mol. Cell Bio.* **2001**, *21*, 1441-1443.

17. Vetter, I. R.; Wittinghofer, A. *Science* **2001**, *294*, 1299-1304.

18. Lu, S.; Jang, H.; Muratcioglu, S.; Gursoy, A.; Keskin, O.; Nussinov, R.; Zhang, J. *Chem. Rev.* **2016**, *116*, 6607-6665.

19. Taparowski, E.; Suard, Y.; Fasano, O.; Shimizu, K.; Goldfarb, M.; Wigler, M. *Nature* **1982**, *300*, 762-765.

20. Reddy, E. P.; Reynolds, R. K.; Santos, E.; Barbacid, M. *Nature* **1982**, *300*, 149-152.

21. Tabin, C.; Bradley, S. M.; Bargmann, C. I.; Weinberg, R. A.; Papageorge, A. G.; Scolnick, E. M.; Lowy, D. R.; Chang, E. H. *Nature* **1982**, *300*, 143-149.

22. Gibbs, J. B.; Sigal, I. S.; Poe, M.; Sconick, E. M. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 5704-5708.

23. Scheffzek, K.; Ahmadian, M. R.; Kabsch, W.; Wiesmüller, L.; Lautwein, A.; Schmitz, F.; Wittinghofer, A. *Science* **1997**, *277*, 333-339.

24. Frech, M.; Darden, T. A.; Pedersen, L. G.; Foley, C. K.; Charifson, P. S.; Anderson, M. W.; Wittinghofer, A. *Biochem.* **1994**, *33*, 3237-3244.

25. Spiegel, J.; Cromm, P. M.; Zimmerman, G.; Grossmann, T. N.; Waldmann, H. *Nat. Chem. Biol.* **2014**, *10*, 613-622.

26. Gysin, S.; Salt, M.; Young, A.; McCormick, F. *Genes & Cancer* **2011**, *2*, 359-372.

27. Holderfield, M.; Nagel, T. E.; Stuart, D. D. *Brit. J. Cancer* **2014**, *111*, 640-645.

28. Chong, C. R.; Jänne, P. A. *Nat. Med.* **2013**, *11*, 1389-1400.

29. Berndt, N.; Hamilton, A. D.; Sebti, S. M. *Nat. Rev. Cancer* **2011**, *11*, 775-791.

30. For one such example: Rao, S.; Cunningham, D.; de Gramont, A.; Scheithauer, W.; Smakal, M.; Humblet, Y.; Kourteva, G.; Iveson, T.; Andre, T.; Dostalova, J.; Illes, A.; Belly, R.; Perez-Ruixo, J. J.; Park, Y. C.; Palmer, P. A. *J. Clin. Oncol.* **2004**, *22*, 3950-3957.

31. Downward, J. *Clin. Cancer Res.* **2015**, *21*, 1802-1809.

32. Wang, Y.; Kaiser, C. E.; Frett, B.; Li, H.-Y. *J. Med. Chem.* **2013**, *56*, 5219-5230.

33. Lin, J.; Wang, W.; Fang, G. *Annu. Rev. Toxicol.* **2014**, *54*, 435-456.

34. Cox, A. D.; Fesik, S. W.; Kimmelman, A. C.; Luo, J.; Der, C. J. *Nat. Rev. Drug Disc.* **2014**, *13*, 828-851.

35. Maurer, T.; Garrenton, L. S.; Oh, A.; Pitts, K.; Anderson, D. J.; Skelton, N. J.; Fauber, B. P.; Pan, B.; Malek, S.; Stokoe, D.; Ludlam, M. J. C.; Bowman, K. K.; Wu, J.; Giannetti, A. M.; Starovasnik, M. A.; Mellman, I.; Jackson, P. K.; Rudolph, J.; Wang, W.; Fang, G. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5299-5304.

36. Sun, Q.; Burke, J. P.; Phan, J.; Burns, M. C.; Olejniczak, E. T.; Waterson, A. G.; Lee, T.; Rossanese, O. W.; Fesik, S. W. *Angew. Chem. Int. Ed. Eng.* **2012**, *51*, 6140-6143.

37. Shima, F.; Yoshikawa, Y.; Ye, M.; Araki, M.; Matsumoto, S.; Liao, J.; Hu, L.; Sugimoto, T.; Ijiri, Y.; Takeda, A.; Nishiyama, Y.; Sato, C.; Muraoka, S.; Tamura, A.; Osoda, T.; Tsuda, K.; Miyakawa, T.; Fukunishi, H.; Shimada, J.; Kumasaka, T.; Yamamoto, M.; Kataoka, T. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 8182-8187.

38. Herrmann, C.; Block, C.; Geisen, C.; Haas, K.; Weber, C.; Winde, G.; Möröy, T.; Müller, O. *Oncogene* **1998**, *17*, 1769-1776.

39. Ostrem, J. M.; Peters, U.; Sos, M. L.; Wells, J. A.; Shokat, K. M. *Nature* **2013**, *503*, 548-551.

40. Lim, S. M.; Westover, K. D.; Ficarro, S. B.; Harrison, R. A.; Choi, H. G.; Pacold, M. E.; Carrasco, M.; Hunter, J.; Kim, N. D.; Xie, T.; Sim, T.; Jänne, P. A.; Meyerson, M.; Marto, J. A.; Engen, J. R.; Gray, N. S. *Angew. Chem. Int. Ed. Eng.* **2014**, *53*, 199-204.

41. Nevola, L.; Giralt, E. *Chem. Commun.* **2015**, *51*, 3302-3315.

42. Patgiri, A.; Yadav, K. K.; Arora, P. S.; Bar-Sagi, D. *Nat. Chem. Biol.* **2011**, *7*, 585-587.

43. Patgiri, A.; Jochim, A.; Arora, P. S. *Acc. Chem. Res.* **2008**, *31*, 1289-1300.

44. Leschiner, E. S.; Parkhitko, A.; Bird, G. H.; Luccarelli, J.; Bellairs, J. A.; Escudero, S.; Opoku-Nsiah, K.; Godes, M.; Perrimon, N.; Walensky, L. D. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 1761-1766.

45. Schafmeister, C. E.; Po, J.; Verdine, G. L. *J. Am. Chem. Soc.* **2000**, *122*, 5891-5892.

46. Mcgee, J. H.; Shim, S. Y.; Lee, S.-J.; Swanson, P. K.; Jiang, Y.; Durney, M. A.; Verdine, G. L. *submitted* **2016**.

47. Boder, E. T.; Wittrup, K. D. *Nat. Biotech.* **1997**, *15*, 553-557.

48. Eckert, D. M.; Malashkevich, V.N.; Hong, L. H.; Carr, P. A.; Kim, P. S. *Cell* **1999**, *99*, 103-115.

49. Funke, S. A.; Willbold, D. *Mol. BioSyst.* **2009**, *5*, 783-786.

50. Liu, M.; Pazgier, M.; Li, C.; Yuan, W.; Li, C.; Lu, W. *Angew. Chem. Int. Ed. Eng.* **2010**, *49*, 3649-3652.

51. Lin, M.; Li, C.; Pazgier, M.; Li, C.; Mao, Y.; Lv, Y.; Gu, B.; Wei, G.; Yuan, W.; Zhan, C.; Lu, W.-Y.; Lu, W. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14321-14326.

52. Zhan, C.; Zhao, L.; Wei, X.; Wu, X.; Chen, X.; Yuan, W.; Lu, W.-Y.; Pazgier, M.; Lu, W. *J. Med. Chem.* **2012**, *55*, 6237-6241.

53. Rabideau, A. E.; Liao, X.; Pentelute, B. L. *Chem. Sci.* **2015**, *6*, 648-653.

54. Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.

55. Mitchell, A. R. *Pept. Sci.* **2008**, *90*, 175-188.

56. Hackeng, T. M.; Rosing, J.; Spronk, H. M. H.; Vermeer, C. *Prot. Sci.* **2001**, *10*, 864-870.

57. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. **1994**, *266*, 776-779.

58. Yan, L. Z.; Dawson, P. E. *J. Am. Chem. Soc.* **2001**, *123*, 526-533.

59. Pentelute, B. L.; Kent, S. B. H. *Org. Lett.* **2007**, *9*, 687-690

60. Wan, Q.; Danishefsky, S. J. *Angew. Chem. Int. Ed.* **2007**, *46*, 9248-9252.

61. Dawson, P. E. *Isr. J. Chem.* **2011***, 51*, 862-867.

62. Bode, J. W.; Fox, R. M. Baucom, K. D. *Angew Chem. Int. Ed. Eng.* **2006**, *45*, 1248-1252.

63. Wu, J.; Ruiz-Rodríguez, J.; Comstock, J. M.; Dong, J. Z.; Bode, J. W. *Chem. Sci.* **2011**, *2*, 1976-1979.

64. Aimoto, S. *Pept. Sci.* **1999**, *51*, 247-265.

65. Hojo, H.; Aimoto, S. *Bull. Chem. Soc. Jpn.* **1992**, *65*, 3055-3063.

66. Hojo, H.; Matsumoto, Y.; Nakahara, Y.; Ito, E.; Suzuki, Y.; Suzuki, M.; Suzuki, A.; Nakahara, Y. *J. Am. Chem. Soc.* **2005**, *127*, 13720-13725.

67. Chen, G.; Wan, Q.; Tan, Z.; Kan, C.; Hua, Z.; Ranganathan, K.; Danishefsky, S. J. *Angew. Chem. Int. Ed. Eng.* **2007**, *46*, 7383-7387.

68. Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* **2000**, *2*, 1939-1941.

69. Wu, B.; Chen, J.; Warren, J. D.; Chen, G.; Hua, Z.; Danishefsky, S. J. *Angew. Chem. Int. Ed.* **2006**, *45*, 4116-4125.

70. Zhang, Y.; Xu, C.; Lam, H. Y.; Lee, C. L.; Li, X. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6657-6662.

71. Tam, J. P.; Miao, Z. *J. Am. Chem. Soc.* **1999**, *121*, 9013-9022.

72. Li, X.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2008**, *130*, 5446-5448.

73. Li, X.; Yuan, Y.; Kan, C.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2008**, *130*, 13225-13227.

74. Yuan, Y.; Zhu, J.; Li, X.; Wu, X.; Danishefsky, S. J. *Tetrahedron Lett.* **2009**, *50*, 2329-2333.

75. Rao, Y.; Li, X.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2009**, *131*, 12924-12926.

76. Wang, T.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2012**, *134*, 13244-13247.

77. Shangguan, N.; Katukojvala, S.; Greenberg, R.; Williams, L. J. *J. Am. Chem. Soc.* **2003**, *125*, 7754-7755.

78. Talan, R. S.; Sanki, A. K.; Sucheck, S. J. *Carbohydr. Res.* **2009**, *344*, 2048-2050.

79. Crich, D.; Sasaki, K. *Org. Lett.* **2009**, *11*, 3514-3517.

80. Wu, X.; Stockdill, J. L.; Wang, P.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2010**, *132*, 4098-4100.

81. Wang, P.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2010**, *132*, 17045-17051.

82. Roberts, A. G.; Johnston, E. V.; Shieh, J.-H.; Sondey, J. P.; Hendrickson, R. C.; Moore, M. A. S.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2015**, *137*, 13167-13175.

Chapter 4: Total Chemical Synthesis and Folding of All-L and All-D KRas(G12V)

Retrosynthesis of KRas(G12V)[1-166]:

Our goal is to identify all-D peptide inhibitors of oncogenic KRas using mirror-image yeast surface display (YSD). Critical to this aim is completion of a total synthesis of the target protein using all-D amino acid residues. Highlighted in **Figure 1** is the primary structure of the KRas G-region, consisting of 166 amino acid residues. We are targeting a G12V mutant commonly associated with various cancers. In our study, we opted to not include the hypervariable region (HVR) as it is not critical for GTPase activity and effector binding.[1] However, it has been recently hypothesized that the HVR of KRas-4B may play a more complex role in its higher association with human cancers than other isoforms due to its ability to bind calmodulin.[2,3] It has also been recently suggested that the HVR of KRas-4B may in fact play a more important structural role in Ras-effector binding events than previously thought.[4,5]

In a retrosynthetic fashion, we proposed that the G-region of KRas could be dissected into five simpler peptidyl sequences obtained through Fmoc-based SPPS. Utilization of the native cysteine residues within the full sequence reduces the synthetic challenge to three NCLs at Cys51, Cys80, and Cys118. We found the synthesis of KRas[118-166] entirely by SPPS to be possible, but inefficient, with poor isolated recovery of the desired peptide. Therefore, this 49-residue sequence was further dissected to enable its efficient preparation. In the absence of available cysteine residues within KRas[118-166] for NCL, two alternative disconnection strategies were hypothesized. One potential option would entail disconnection between Ser145-Ala146,

in which a non-native Cys146 residue would be utilized for NCL followed by MFD (Cys146 → Ala146) to yield the native sequence.[6] Alternatively, the chemoselective isonitrile-mediated activation of C-terminal thiocarboxylic acids represents a complementary ligation strategy that does not rely on cysteine. Given recent precedent from our laboratory, we hypothesized that a C-terminal thioacid at Gly138 could serve as a potential disconnection site.[7] Here, Gly138 was chosen as a median residue site and selected as an ideal acyl donor that cannot epimerize upon activation. Due to associated sequence-specific challenges, this retrosynthetic disconnection would provide a worthwhile opportunity to further demonstrate the potential of this ligation strategy. Notably, this ligation (Gly138-Ile139) would require the efficient bimolecular reaction of a hindered β-branched nucleophile at Ile139. Additionally, this strategy would require efficient orthogonal protection and deprotection of side-chain nucleophiles at Lys128, Lys147, and Lys165.

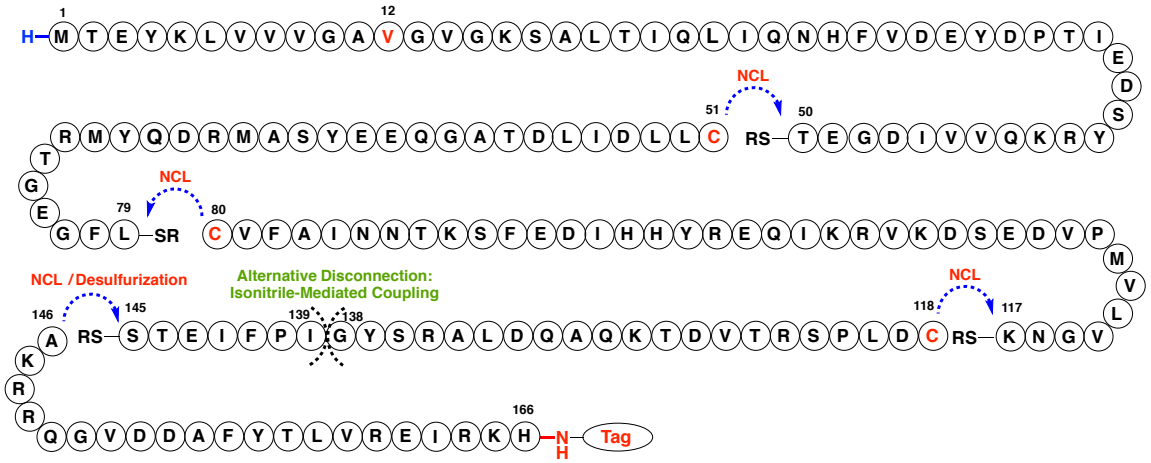
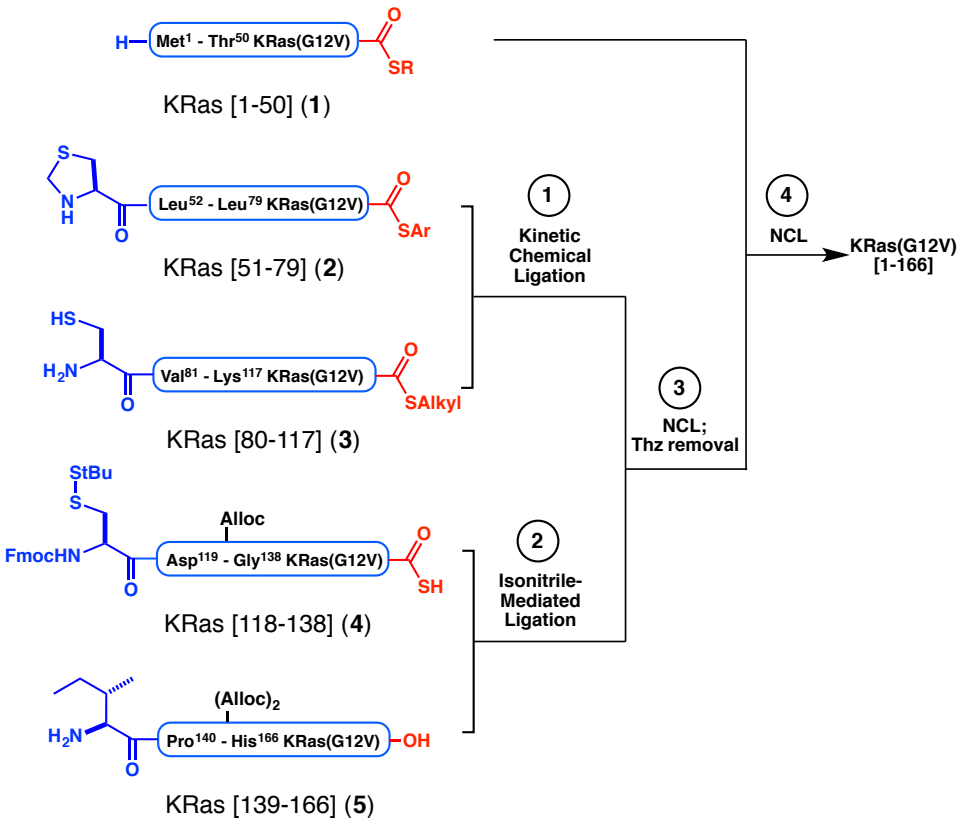


Figure 1: Retrosynthetic disconnection sites toward KRas(G12V) [1-166]

Shown in **Scheme 1** is our first generation approach towards KRas(G12V)[1-166]. Projected steps would include a kinetically controlled NCL between arylthioester **2** and alkyl thioester **3** would provide KRas[51-117].[8] Coupling of KRas[118-138] (**4**)

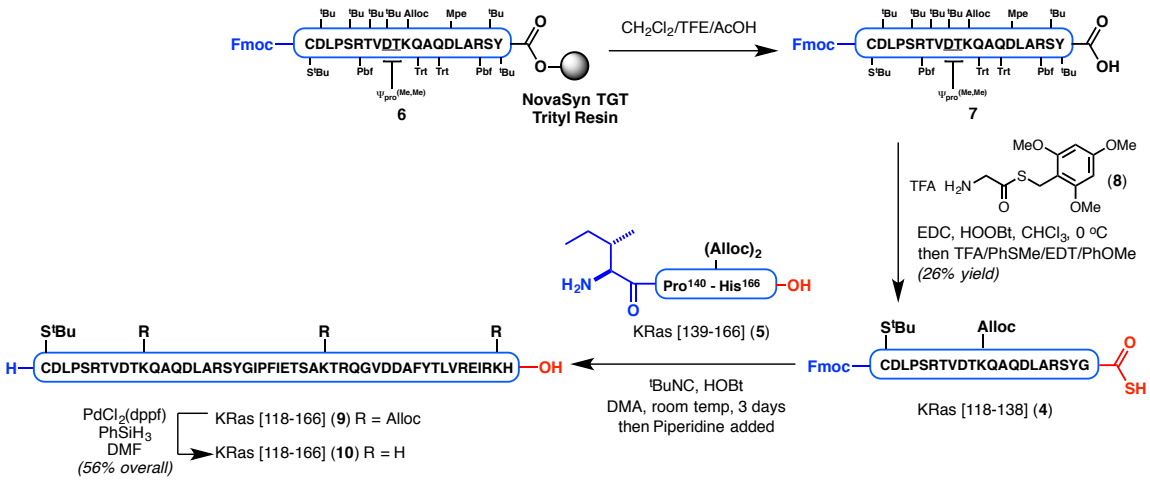
and KRas[139-166] (**5**) would be accomplished using an isonitrile-mediated ligation, or alternatively assembled via NCL-MFD as described. The merger of KRas[51-117] with KRas[118-166] by NCL followed by removal of the *N*-terminal thiazolidine (Thz) protecting group would then yield KRas[51-166].[9] Finally, NCL of KRas[1-50] (**1**) with KRas[51-166] would yield the desired sequence. This strategy was designed with the intention of synthesizing a variety of oncogenic mutated KRas isoforms, with mutations commonly at Gly12 or Gly13 within the P-loop. Therefore, sequence mutants of KRas[1-50] could be quickly synthesized and convergently ligated to KRas[51-166] in a final step.



Scheme 1: First generation approach toward KRas(G12V)

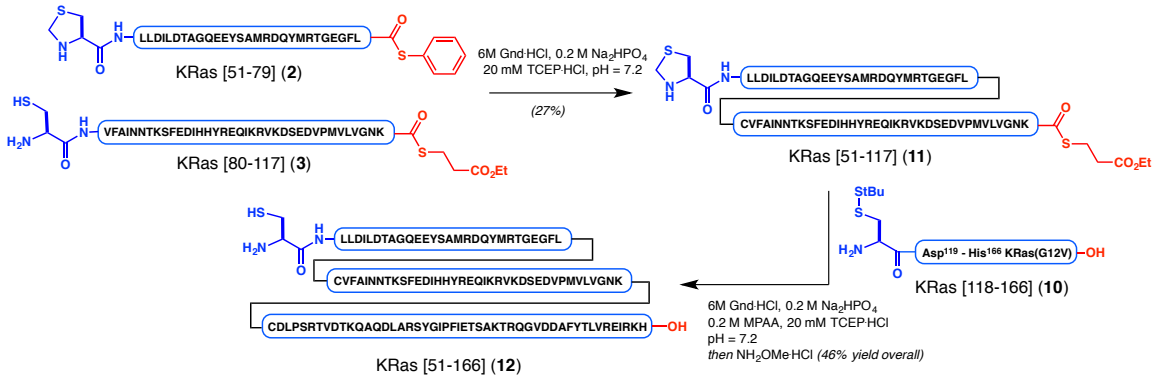
Synthesis of KRas(G12V)[1-166]

In the forward direction, the synthesis of KRas[118-166] was indeed accomplished utilizing an isonitrile-mediated ligation approach (**Scheme 2**). For controlled ligation between **4** and **5**, orthogonal protection of the three Lys side-chain ε-amines as well as the C-terminal Cys β-thiol group were necessary. Lys side-chains were protected as allyloxycarbamates (Alloc). Cys118 was orthogonally protected as a *tert*-butyl disulfide. The C-terminal thioacid was synthesized by Sakakibara C-terminal elongation of fully protected peptide **7**, utilizing H-Gly-STmob (**8**).[10] Upon acid mediated global side-chain deprotection, the C-terminal thioacid is unveiled to yield **4** in 26% yield. Upon activation with *tert*-butyl isocyanide in the presence of HOBt, **4** and **5** undergo coupling to efficiently generate the Alloc-protected ligation product. It was found that the direct addition of piperidine effects *N*-terminal Fmoc removal in the same reaction pot. This crude intermediate could be utilized in a subsequent Alloc-removal step under reductive conditions (Pd, $PhSiH_3$). Final HPLC purification afforded the fully-deprotected fragment **10** in 56% yield.



Scheme 2: Synthesis of KRas[118-166] (**10**) via isonitrile-mediated activation of **4.**

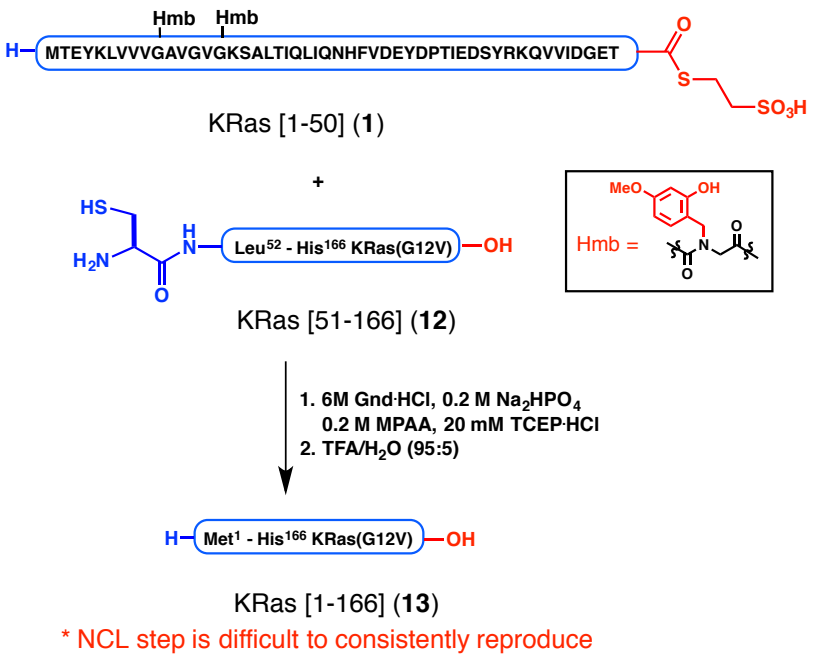
With an efficient synthesis of KRas[118-166] **10** at hand, we turned our attention to completion of the full length sequence. Aryl thioester KRas[51-79] **2** and alkyl thioester KRas [80-117] **3** were ligated under kinetically-controlled NCL conditions to selectively deliver **11** in 27% yield. This sequence was subjected to ligation with KRas[118-166] (**10**) via NCL, followed by *in situ* removal of the N-terminal Thz protecting group upon the addition of methoxylamine hydrochloride. This two-step one pot process provided **12** isolated in 46% yield.



Scheme 3: Synthesis of KRas[51-166] (**12**)

With **12** in hand, we attempted ligations with KRas[1-50] (**1**). We found the synthesis and handling of this sequence to be extremely difficult, likely due to the hydrophobic nature of the N-terminal sequence. Full-length Ras is known to be hydrophobic and insoluble, especially in the absence of a bound nucleotide substrate.[11] Becker and co-workers completed a synthesis of HRas via the ligation of HRas[1-50] with HRas[51-166].[12] However, they note similar solubility problems and required an amphiphilic detergent additive in this ligation. Moreover, the final protein was obtained with a loss of 18 Da, suggesting an unassigned dehydration or deamination took place during the ligation or during purification. In our hands, synthesis of **1** was greatly aided

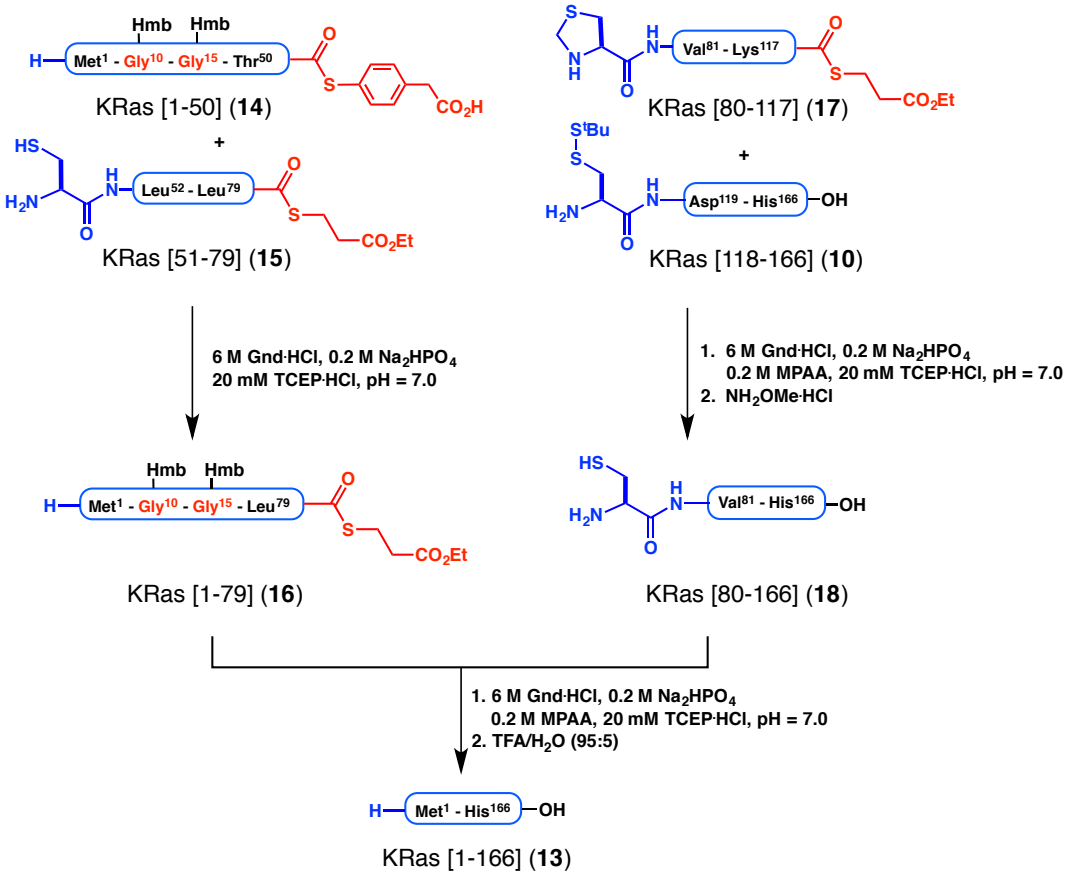
by strategic placement of acid labile 2-hydroxy-4-methoxybenzyl (Hmb) groups incorporated within the amide backbone.[13] Installation of a tertiary amide within peptides is known to disrupt secondary structure and minimize aggregation. This in turn improves SPPS efficiency and augments the solubility of difficult sequences.[14,15,16] By conventional design, the Hmb group can also be rendered temporarily acid-stable if the phenol moiety is acylated. We found that Hmb groups improved synthesis of **1**, which had relatively high solubility in guanidine buffer and could be directly applied in NCL with **12** without the need for solubilizing agents. Furthermore, the ligation product **13** could be deprotected upon treatment with TFA/H_2O without dehydration. However, the yield and ease of purification for this NCL was difficult to consistently reproduce. The reason for this inconsistency is still unclear. We noted that different batches of peptide **1** tended to undergo NCL with varying efficiency. In some cases, aggregated peptide would precipitate from the ligation mixture, wherein the desired protein could not be recovered. Moreover, in successful cases, HPLC separation of product from unreacted starting materials was difficult, as their HPLC retention times are coincident.



Scheme 4: Capricious NCL between **1** and **12** followed by Hmb removal

In an effort to improve the reproducibility of our synthetic route, it was hypothesized that altering the ligation order of events to give full-length protein may address solubility and purification issues (**Scheme 5**). Perhaps by initially ligating KRas[1-50] (**14**) with KRas[51-79] (**15**), issues in aggregation previously observed could be attenuated in the final steps. This strategy would require only slight modifications in N-terminal protecting groups and C-terminal thioesters. Indeed, **14** successfully underwent kinetic chemical ligation with **15** to afford KRas[1-79] **16** in 36% isolated yield. NCL between **17** and **10** proceeded smoothly, yielding KRas[80-166] **18** in 56% yield after Thz removal. Finally, NCL at this new disconnection point between Leu79 and Cys80 proceeded to yield the full-length sequence. Following treatment with TFA/H_2O, **13** was obtained in 19% yield over 2 steps. Moreover,

unreacted starting materials **16** and **18** are easily separated by HPLC purification, with greatly different retention times from the full-length sequence.



Scheme 5: Revised successful synthetic route towards KRas(G12V) [1-166] (**13**)

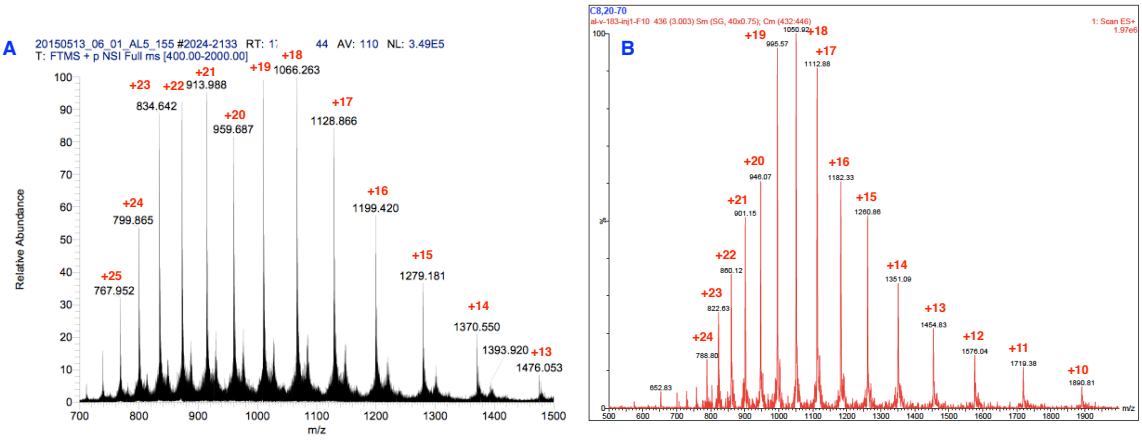


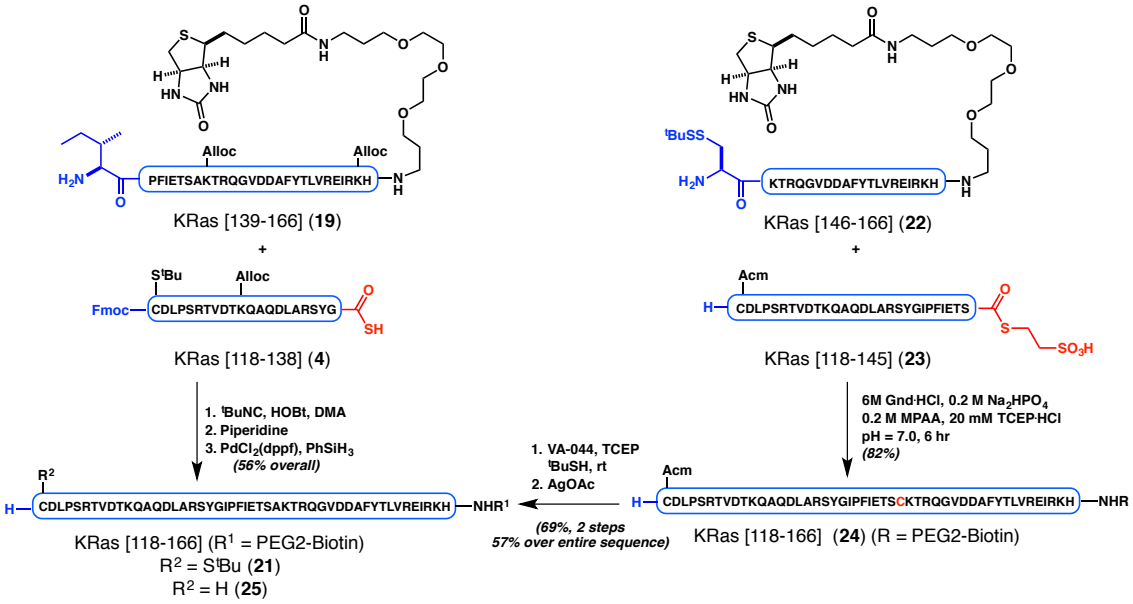
Figure 2: (**A**) High resolution mass spectrum of Hmb-protected KRas[1-166]; (**B**) ESI-(+)-MS of deprotected KRas[1-166] (**13**)

Application of Optimized Synthetic Route to the Synthesis of a Biotinylated KRas(G12V) Variant

With a reproducible synthetic route at hand, this chemistry was replicated to incorporate a biochemical label within KRas(G12V). We chose to label the C-terminus of the peptide, because it is furthest from nucleotide and effector binding sites. Two new C-terminal fragments, **19** and **22**, were synthesized by SPPS using an Fmoc-PEG Biotin NovaTag™ resin. These two fragments were then applied in a comparative synthesis of biotin-labeled KRas[118-166], featuring two projected synthetic routes utilizing chemistry developed in our laboratory (**Scheme 6**). In one permutation of our synthesis, **19** efficiently underwent isonitrile-mediated ligation with **4** similarly as previously described. Shown in **Figure 3** is characterization of each step in this sequence by UPLC. After 44 hours of reaction (**Figure 3A**), only a trace amount of nucleophilic peptide fragment **19** remains. Upon addition of piperidine (**Figure 3B**), Fmoc removal is efficient. Interestingly, in addition to observed hydrolysis at the C-terminus of KRas[118-138], the reaction of the C-terminus with piperidine was also observed, yielding the corresponding C-terminal amide. This result suggests that even after 44 hours, there is still active acyl donor present in the reaction mixture. Finally, after Alloc removal, purified **21** (**Figure 3D**) is obtained in 56% isolated yield over the 3-step sequence.

We compared this route to a more conventional NCL-MFD strategy. Utilizing a non-native C146 residue, **22** underwent NCL with thioester **23** as anticipated to yield **24** in 82% isolated yield. This fragment was subjected to radical dethiylation upon treatment with VA-044, TCEP, and *tert*-butylthiol with high effiency at room

temperature. Finally, Acm removal upon treatment with AgOAc provided the analogous fragment KRas[118-166] (**25**). Overall, this 3-step sequence afforded **25** in 57% yield. The comparable ease of reaction setup in each study suggests that in certain cases, isonitrile-mediated ligations at C-terminal Gly thioacids can also provide facile access to large peptides. This method should be considered as a complementary strategy to NCL-MFD, and as an alternative option especially for hydrophobic sequences lacking available or median cysteine residues. Moreover, although both of these synthetic strategies require multi-step assembly and deprotection, we contend that the isonitrile-mediated ligation approach may be preferable due to intermediary purification ease in cases where the overall yields are similar – requiring only precipitation upon addition of diethyl ether with no desalting or HPLC necessary to isolate intermediates.



Scheme 6: Comparative syntheses of biotinylated KRas[118-166] using isonitrile-mediated ligation or NCL-MFD strategies.

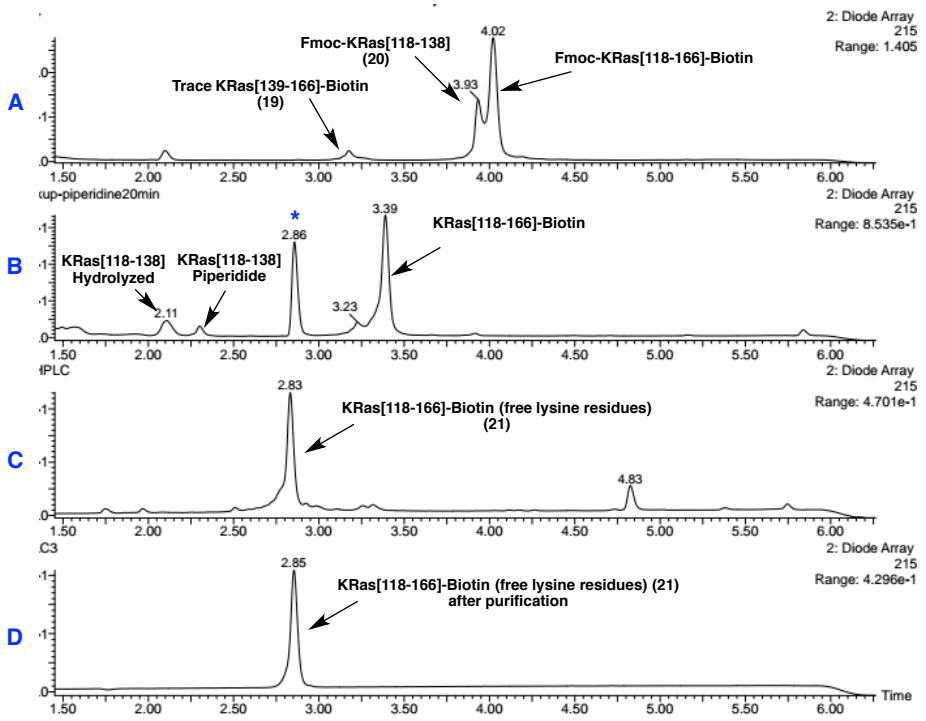
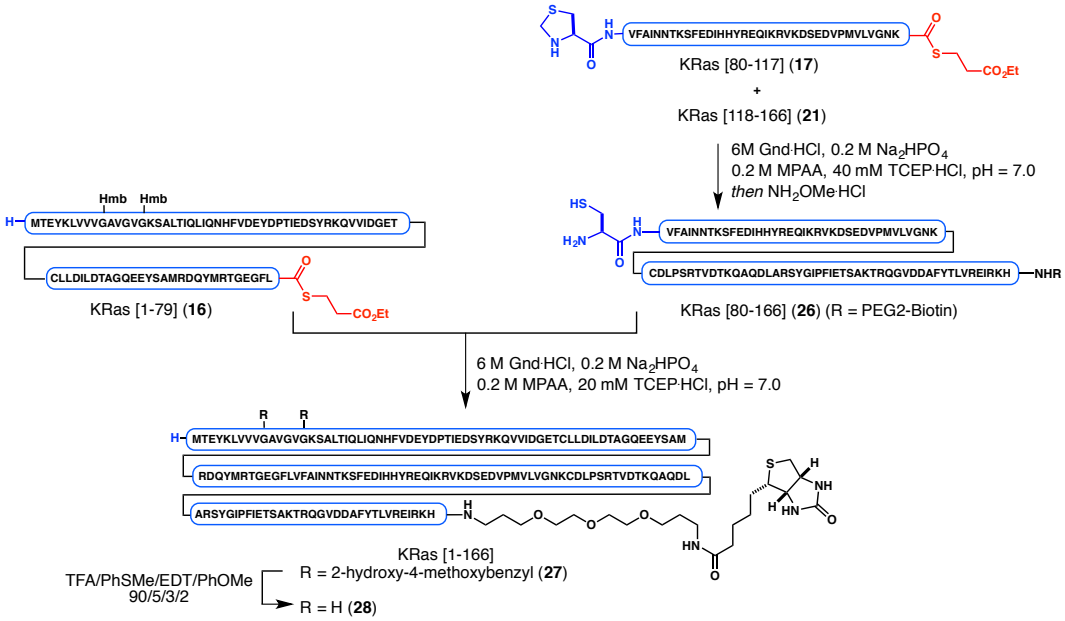


Figure 3: UPLC traces for the isonitrile-mediated coupling sequence between **19** and **4**. (**A**) Isonitrile-mediated ligation (t = 44 hr); (**B**) Fmoc removal upon addition of piperidine (t = 30 min; *denotes non-peptidyl UV peak, which disappears after ether precipitation, and is assumed to be derived from piperidine or Fmoc); (**C**) Crude alloc removal; (**D**) Purified KRas[118-166] (**21**).

Completion of biotinylated KRas(G12V) was achieved as outlined in **Scheme 7**. Peptides **17** and **21** underwent NCL as anticipated, followed by Thz removal to afford **26** in 63% isolated yield. Finally, NCL with KRas[1-79] (**16**) efficiently afforded the desired Hmb-derived full-length sequence, with full consumption of **16** observed within 6-8 hours at neutral pH. Moreover, it was found that the ligation product (**27**) could be nearly entirely recovered by exploiting solubility differences between **26** and **27**. Upon dilution of the NCL with H_2O, **26** remains entirely dissolved while the hydrophobic ligation product **27** precipitates and can be collected by centrifugation (**Figure 4**). Exposure of the resulting precipitate to Cocktail R effects quantitative Hmb removal with suppression of methionine oxidation to afford biotinylated KRas(G12V)[1-166] in

21% yield over the entire sequence. Considering reactants fully converge to the desired product during each step in this sequence, the relatively low yield is largely attributed to the hydrophobic nature of full-length KRas, which complicates solubility during purification.



Scheme 7: Synthesis of biotinylated KRas(G12V)[1-166] (**28**)

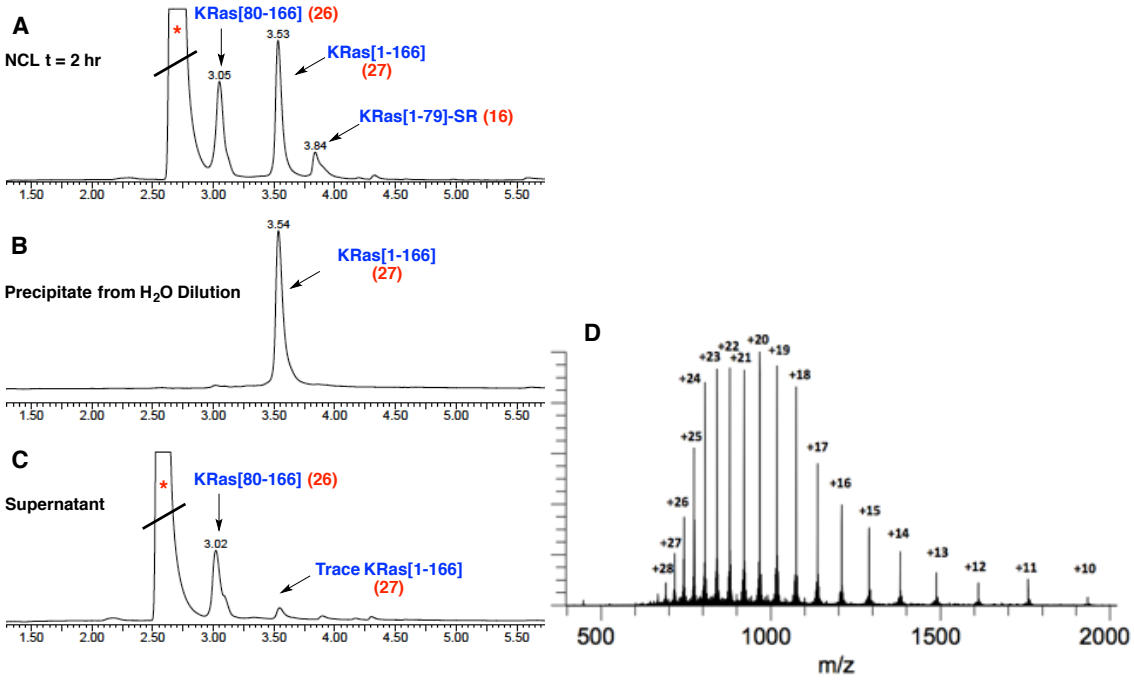
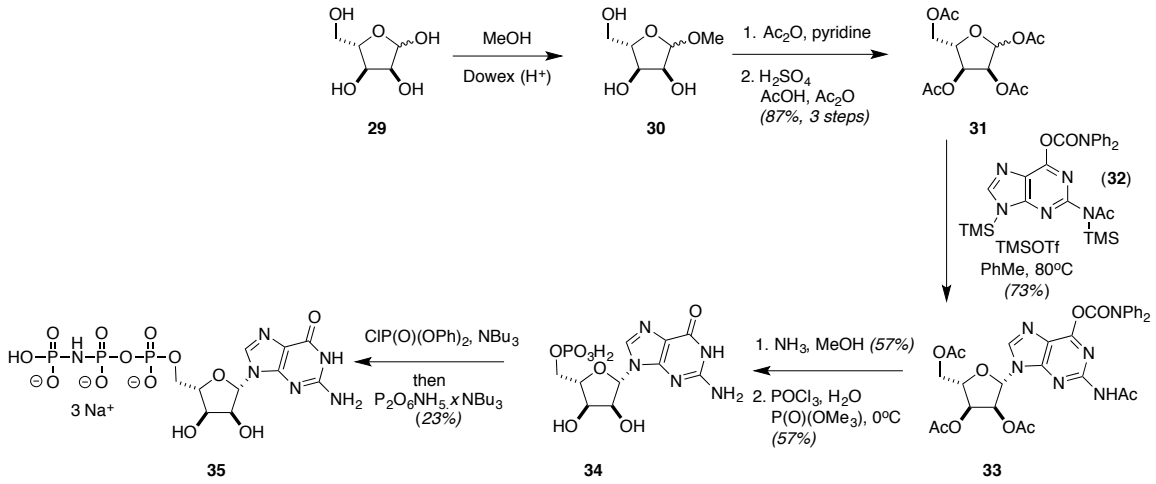


Figure 4: (**A**) NCL between **16** and **26** (t = 2 hr). * denotes absorption peak from MPAA (**B**) Isolated **27** from H_2O dilution of the NCL (after completion, t ~ 6.5 hr) (**C**) Supernatant from H_2O dilution, containing excess **27**. Only trace **16** remains solubilized in the supernatant. (**D**) High resolution mass spectrum of deprotected **28** following treatment with Cocktail R.

With all-L and all-D variants of biotinylated KRas(G12V) in hand using the route described in **Scheme 7**, we sought to establish the biochemical competency of this enantiomeric pair of proteins in collaboration with the laboratory of Gregory Verdine at Harvard University. Folding conditions for the all-D protein were established using the mirror-image L-nucleotide as substrate, synthesized according to literature protocols. Highlighted in **Scheme 8** is our synthetic route towards L-GppNHp as a non-hydrolyzable GTP analog. Starting with L-ribose (**29**), treatment with methanol under acidic conditions generated kinetically-favored methyl glycoside **30**. Subsequent acetylation and exchange of the methyl glycoside for acetate afforded L-ribose pentaacetate **31** in 87% yield over the entire sequence.[17] Regio- and stereoselective glycosylation with protected guanine derivative **32** afforded protected guanosine **33**.[18] Upon global deprotection with ammonia, L-Guanosine was subjected to regioselective phosphorylation using Nishizawa's conditions to afford the corresponding 5'-monophosphate (**34**).[19] Finally, utilizing a procedure described by Yount entailing activation of the monophosphate with diphenylphosphoryl chloride, followed by displacement with tributylammonium imidodiphosphate, **34** was converted to enantiomeric GppNHp **35** and isolated as the sodium salt.[20]



Scheme 8: Synthesis of L-GppNHp (**35**).

We adopted a previously reported refolding protocol, which entailed solubilization of the proteins in guanidine buffer and rapid refolding into denaturant-free buffer, yielding monomeric proteins with moderate efficiency.[12,21] Analytical gel filtration shows an initial eluting peak assigned as aggregate, with a second peak eluting with similar retention time as the standard recombinant KRas(G12V) (**Figure 5**). Monomeric proteins could be isolated by gel filtration with greater than 90% purity, and CD analysis of the folded proteins closely matched that of recombinant KRas. All-D residue KRas(G12V) properly bears opposite signal to that of synthetic and recombinant L-residue KRas(G12V). Comparison of A_{260}/A_{280} values showed an increase from 0.6-0.7 for the unfolded to 1.0-1.1 for the recombinant and folded proteins, consistent with association with guanine nucleotides, which possess an absorption maximum at 252 nm.[22]

To further probe the nucleotide-binding of our synthetic proteins, we refolded both in the presence of fluorescent mant-GppNHp ((2'/3')-*O*-(*N*-

methylanthaniloyl)guanosine-5'-*O*-[(β,γ)-imidotriphosphate]) analogs, which exhibit an increase in fluorescence upon binding to Ras. This fluorescence output phenomenon has been used to study Ras-ligand binding interactions.[23] The enantiomeric pair of proteins folded with mant-GppNHp nucleotides of corresponding stereochemistry exhibited comparable nucleotide dissociation behavior relative to recombinant KRas. This is marked by a decrease in fluorescence upon the addition of excess unlabeled GppNHp of correlated stereochemistry (**Figure 5D-F)**. Interestingly, preliminary experiments found that at high assay concentration of nucleotide, both synthetic all-D KRas(G12V) and recombinant KRas(G12V) underwent nucleotide exchange with either antipode of nucleotide. However, at lower concentration (10 μM nucleotide), good exchange selectivity was observed only for the nucleotide of correlated stereochemistry.

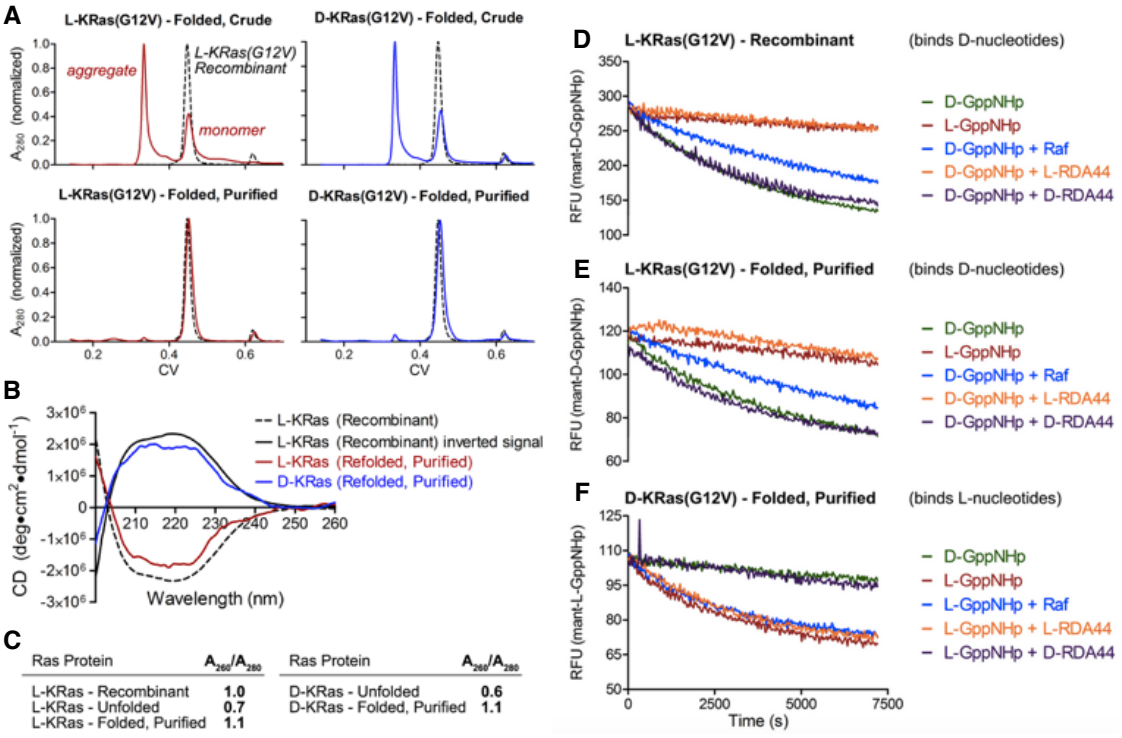


Figure 5: Analysis of folded proteins (**A**) Crude and purified folded synthetic proteins; (**B**) CD analysis of folded synthetic proteins and recombinant KRas(G12V); (**C**) Comparative analysis of A_{260}/A_{280} values for unfolded and folded synthetic proteins; (**D**) Nucleotide exchange assay for recombinant KRas(G12V); (**E**) Nucleotide exchange assay for synthetic L-KRas(G12V); (**F**) Nucleotide exchange assay for synthetic D-KRas(G12V).

We also investigated the effect of adding different Ras-binding peptide ligands to our synthetic proteins. This included addition of the Ras-binding domain (RBD) of B-Raf, a Ras effector known to slow dissociation, as well as RDA44, a recently discovered Ras-binding miniprotein that stops nucleotide dissociation.[24] RDA44 was added either as the all-L residue or all-D residue peptide. In the case of synthetic L-KRas(G12V), the addition of the B-Raf RBD slowed dissociation in a similar manner to that observed with recombinant KRas(G12V). The addition of the B-Raf RBD had no observed effect on nucleotide dissociation from synthetic D-KRas(G12V). Similarly, addition of RDA44 (all L-amino acid residues) slowed nucleotide dissociation for

recombinant and synthetic L-KRas, with no effect observed on all D-residue KRas. Conversely, addition of an all D-residue version of RDA44 does indeed slow nucleotide dissociation for all-D KRas as anticipated and had no effect on recombinant or synthetic L-KRas. These results demonstrate that folding, nucleotide association, and peptide binding profiles of both mirror-images of synthetic KRas(G12V) resemble that of recombinant KRas(G12V) and exhibit the expected enantiodiscrimination.

With the *in vitro* biochemical activity of our synthetic pair of enantiomeric proteins confirmed, future steps will be taken to utilize this material in the search for new cancer therapeutics. We anticipate that our synthetic folded protein will be a useful tool in mirror-image YSD for the discovery of new, all-D peptide inhibitors of Ras-effector interactions. In preparation for these studies, over 5 mg (~260 nmol) of all-D residue KRas(G12V) was recently prepared through this method. Moreover, chemical protein synthesis enables site-specific modification of the target of interest, and our described synthetic route may enable library synthesis of oncogenic forms of all-D KRas for use in mirror-image YSD.

Further Exploration of Isonitrile-Mediated Ligations: Application to a Total Chemical Synthesis of HIV-1 Protease

The HIV-1 protease is an aspartyl protease enzyme essential for the proliferation of the HIV virus which causes AIDS.[25] The protease serves to post-translationally process proteins critical to the HIV virus replication and structure, such as reverse transcriptase, integrase, and structural proteins that make up the viral coat and core. For these reasons, HIV-1 protease inhibitors are an important class of drugs in retroviral treatment.[26]

Structurally, the protease is a 21.4 kDa homodimeric protein, consisting of two 99-residue monomeric polypeptides (**Figure 6**). HIV-1 protease and non-natural analogs thereof have been chemically synthesized through a variety of strategies. These include synthesis of its entire sequence by SPPS as well as modular syntheses via ligation strategies.[27-34] A large body of early chemical synthesis, structural and biochemical studies on this enzyme were conducted in the Kent laboratory.

One of the strategies targeting this protein relies on the use of NCL-desulfurization to ligate smaller fragments synthesized by SPPS. Recently, Qi and co-workers completed a synthesis of azide-labelled HIV-1 protease (I66A), such that an azide-alkyne cycloaddition could be utilized in later studies.[35] Their preliminary strategy toward this goal of bioorthogonally-labeled HIV-1 protease (**Scheme 9**) relied on integration of an alkyne within the sequence via incorporation of Fmoc-propargylglycine (Fmoc-Pra) replacing native Lys41. In this approach, the full-length protein was disconnected at Ala28 and Ala71, with projected NCLs at these positions followed by MFD (Cys → Ala). A 16-residue polyarginine solubility tag containing a portion of immature HIV-1 protease, initially devised by Kent and co-workers, was appended at the C-terminus.[32] This Arg-rich tag abrogates known aggregation within the C-terminal region of the protein and facilitates handling and purification. After assembly and folding, the protease is known to self-cleave this solubility tag through an autoprocessing mechanism to reveal the 99-residue protein as its active homodimer. This autoprocessing is mechanistically similar to maturation of HIV-1 protease *in vivo*.[36] Despite successful NCL disconnections, the propargylglycine residue ultimately

proved unstable to free radical desulfurization, affording a covalent adduct with TCEP (**41**) under these conditions and barring the use of their NCL-desulfurization strategy.

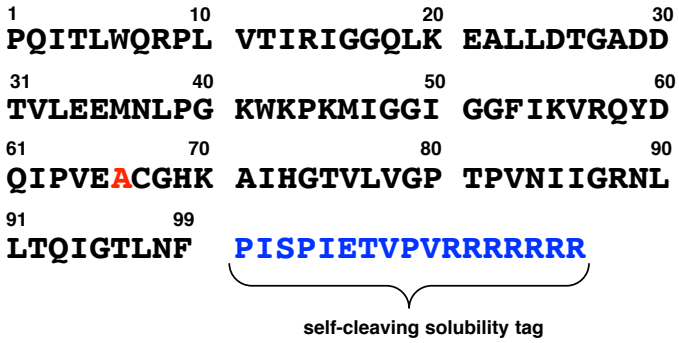
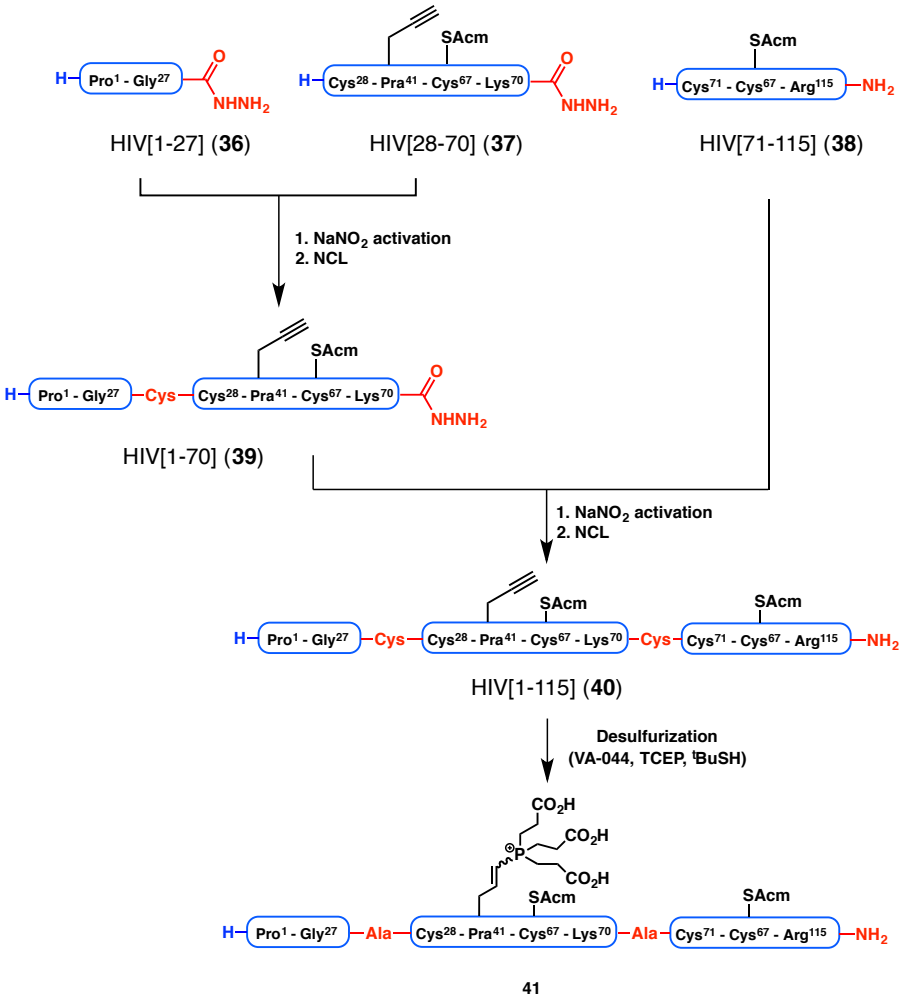


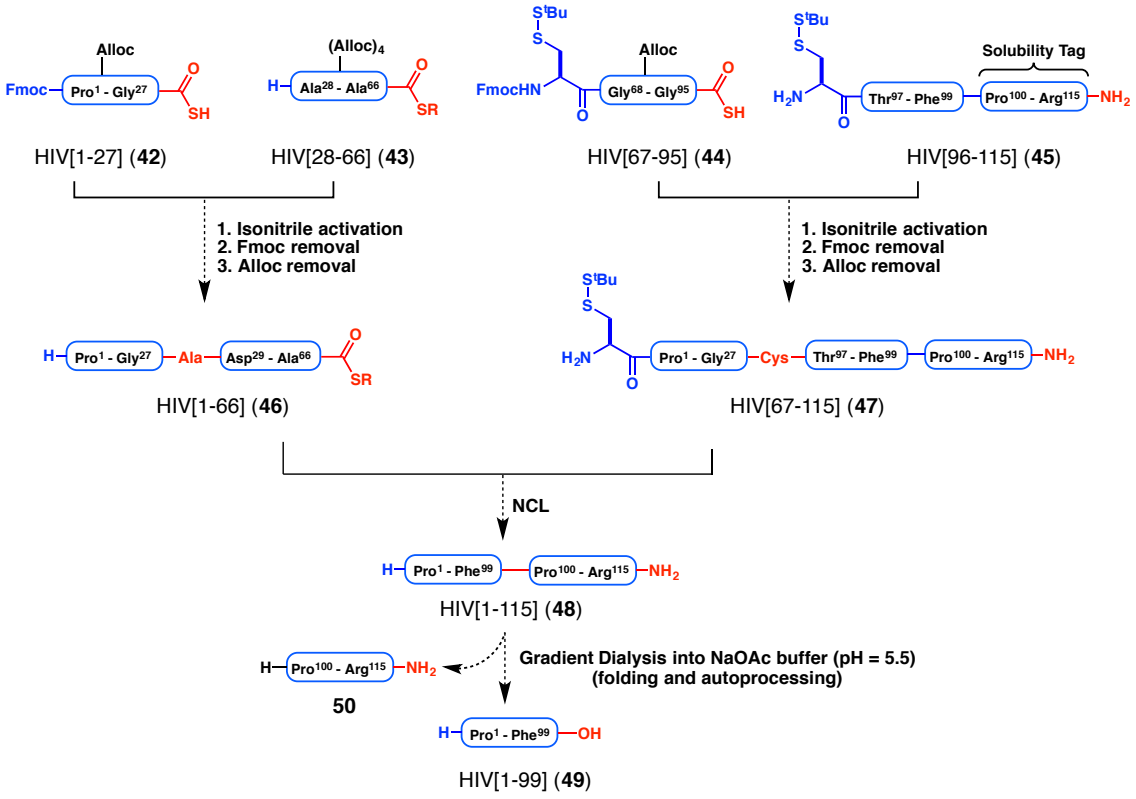
Figure 6: Primary sequence of HIV-1 Protease (I66A) and appended Arg-rich sequence.



Scheme 9: Failed NCL-MFD approach towards alkynylated HIV-1 Protease

With these results indicating that functionality such as alkynes and azides are incompatible with MFD, we wondered if a combined isonitrile-mediated ligation and NCL strategy could serve as a complementary approach toward the synthesis of HIV-1 protease. The I66A mutation does not affect protease activity, and targeting an I66A mutant would reduce the steric impediment of an NCL at that disconnection site.[35] Shown in **Scheme 10** is our projected synthesis toward this target.

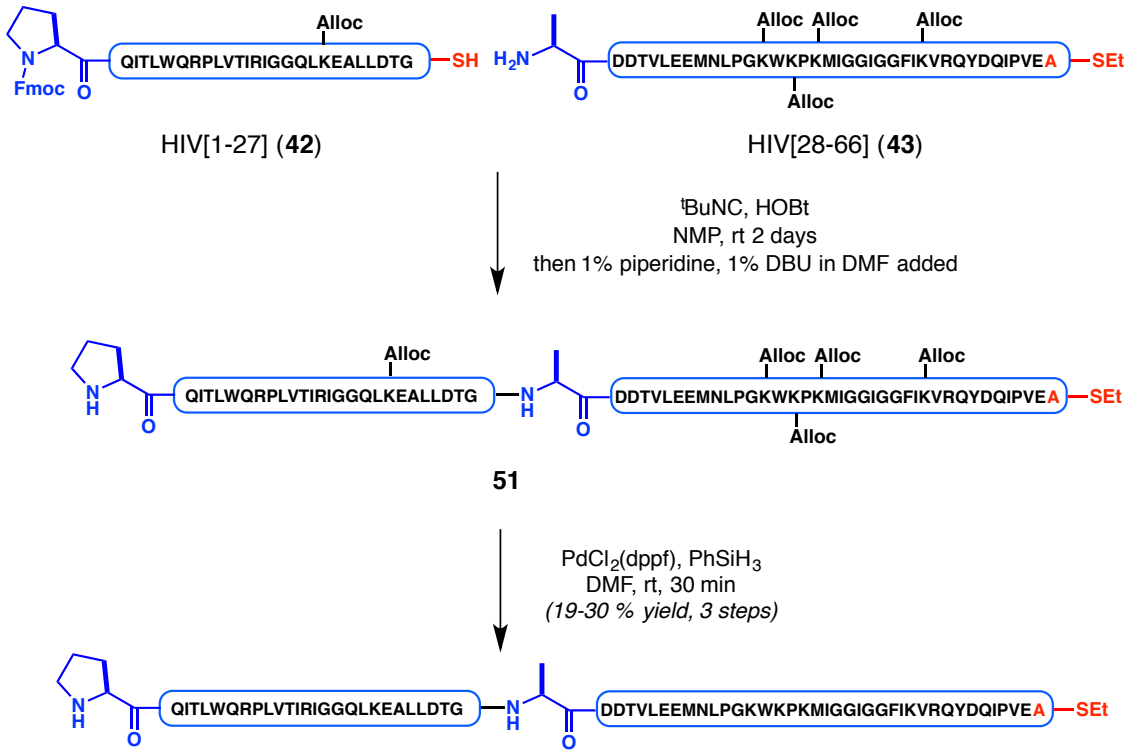
As in earlier examples of isonitrile-mediated ligation described for the synthesis of KRas[118-166], lysine residue side-chains will be orthogonally masked using Alloc groups, and cysteine will be incorporated as the *tert*-butyl disulfide. In this approach, HIV[1-66] (**46**) would be synthesized via the chemoselective isonitrile-mediated ligation between thioacid **42** and thioester **43**. In particular, this disconnection would require the efficient removal of five Alloc-protected lysine sidechains in addition to the N-terminal Fmoc group. Moreover, this approach would require us to demonstrate the compatability of a C-terminal thioester on **43** − functionality that poses conceivable complications during each of the three steps in the ligation protocol. Separately, HIV[67-95] (**44**) would undergo isonitrile-mediated ligation with HIV[96-115] (**45**), which contains a polyarginine solubility tag, to generate **47**. Peptides **46** and **47** would then be subjected to NCL conditions reliant on native Cys67 to generate the full-length sequence, HIV[1-115] (**48**). Finally, **48** would undergo autoproteolytic cleavage to remove the solubilizing polyarginine tag according to literature precedent to afford folded HIV protease [1-99] (**49**). Moreover, successful autoprocessing will serve to further demonstrate that isonitrile-mediated ligations can be used to produce catalytically active enzyme.



Scheme 10: Proposed synthesis of HIV-1 Protease (I66A) using combined isonitrile-mediated ligation and NCL strategies.

In practice, our devised synthetic scheme was successful. Thioacids **42** and **44** could be accessed relatively easily through synthesis. However, difficulty in solubility of these partially protected peptides in aqueous solvents − likely due to loss of cationic charge as a result of amine protection − precluded isolation of large quantities of these materials. Nonetheless, **42** underwent isonitrile activation and chemoselective ligation with thioester **43** (**Scheme 11**). After 48 hours, the reaction was quenched upon the addition of a solution of DBU and DMF, which efficiently removed the N-terminal Fmoc group without affecting the C-terminal thioester. Finally, global removal of five Alloc-protected lysine sidechains was effected upon treatment of this crude intermediate with $PdCl_2$(dppf) and $PhSiH_3$ to afford deprotected ligation product **46** in 19-30% yield.

Again, the C-terminal thioester remained inert under these deprotection conditions, with no observed side reactions such as Fukuyama-type reduction of the alkyl thioester.[37]



Scheme 11: Synthesis of HIV Protease (I66A) [1-66] via isonitrile-mediated ligation

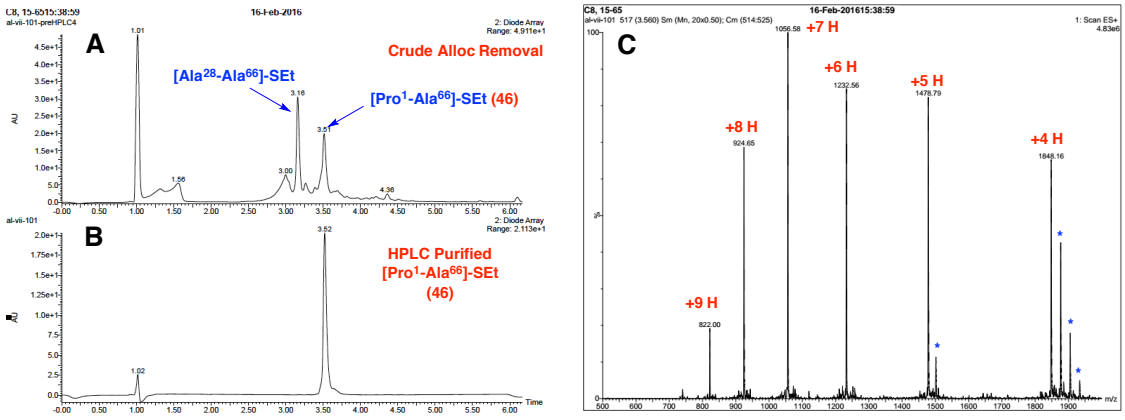
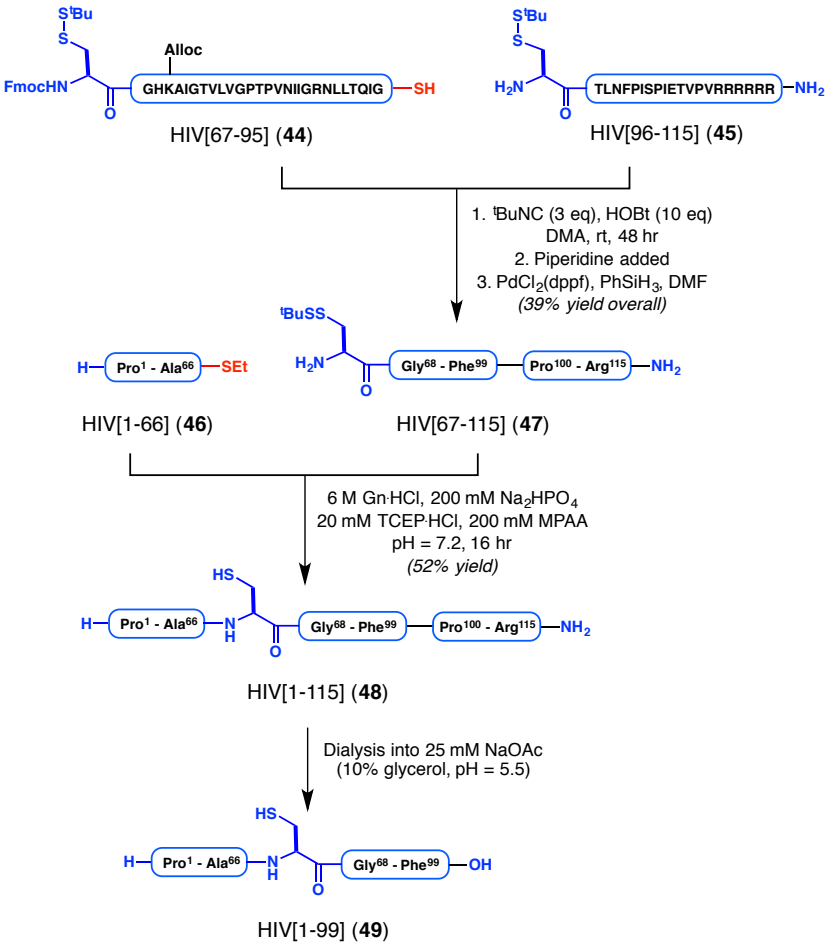


Figure 7: Characterization of isonitrile-mediated coupling sequence between **42** and **43** (**A**) UPLC trace of crude Alloc removal step in the synthesis of **46**; (**B**) UPLC trace of purified **46**; (**C**) Integrated ESI-(+)-MS trace of **46.** * Denotes associated TFA molecules.

Completion of the synthesis of HIV-1 protease was accomplished through an isonitrile-mediated ligation between **44** and **45**, providing **47** in 39% yield over the entire sequence (**Scheme 12**). With the two halves of the protease assembled via isonitrile-mediated ligation, we then took advantage of native Cys67 for NCL between **46** and **47**, providing the full-length 115-residue sequence **48** in 52% isolated yield. With realization of our combined strategy, we were able to verify that our strategy affords proper bond-disconnections and catalytically-active protein by applying a modified literature protocol for cleavage of the C-terminal polyarginine tag. This was accomplished by initial dissolution of **48** in guanidine buffer, followed by gradient dialysis at 4 ^{o}C into 25 mM NaOAc (pH = 5.5, containing 10% glycerol). After 10 hours, only a trace amount of unfolded protein remained solubilized, with the major constituent corresponding to proteolysis between Phe99 and Pro100 (**Figure 8**).



Scheme 12: Completion of the synthesis of HIV-1 Protease and autoproteolytic cleavage upon folding.

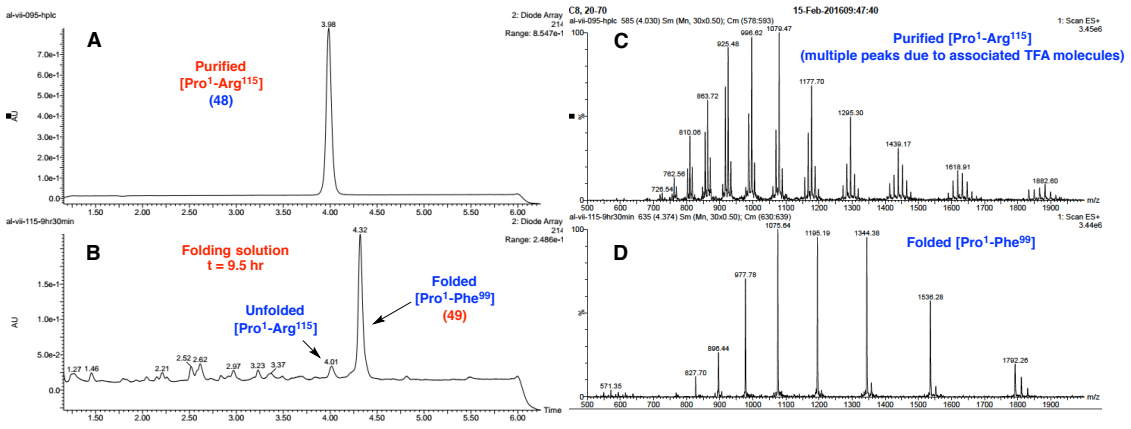


Figure 8: Folding of synthetic HIV-1 protease (**A**) UPLC trace of **48** (λ = 214 nm); (**B**) UPLC trace of folding solution (t = 9.5 hr, λ = 214 nm); (**C**) ESI-(+)-MS of purified **48**; (**D**) Integrated ESI-(+)-MS of folded HIV-1 Protease (**49**).

Future Outlook:

The studies in this section highlight the potentially general utility of the isonitrile-mediated ligation in the synthesis of large polypeptides and proteins. Application of an isonitrile-based approach enabled a total chemical synthesis of KRas(G12V), which displayed biochemical viability and nucleotide binding. Moreover, a comparative analysis of this strategy with a more conventional NCL-MFD strategy in the synthesis of biotin-labeled KRas[118-166] gave similar overall yields for both three-step sequences. We further demonstrated that this method, even without optimization, enabled completion of a total synthesis of HIV-1 Protease (I66A). Importantly, a C-terminal alkyl thioester was compatible with our isonitrile-mediated ligation sequence, enabling a ligation sequence in the N to C direction. These results highlight that isonitrile-mediated peptide ligation should be considered by others in the field of chemical protein synthesis, especially in cases where NCL-MFD is inefficient or incompatible with specific functionality. Moreover, in such an instance, this strategy can be viewed as an alternative to kinetically controlled NCL.[8]

One major limitation to our method thus far is the requirement for orthogonal amine protection of Lys residues and the N-terminus of the peptide fragments. Examples shown in this thesis have used orthogonal protection either in the form of Fmoc or Alloc protecting groups. Although we have demonstrated that even as many as five Alloc-protected amines can be globally deprotected, a major limitation is the hydrophobicity they impart on peptidyl fragments, complicating handling and purification. This is especially true for larger peptide fragments, which are often inherently hydrophobic. Nakahara and co-workers have previously described this issue

in the context of Ag(I)-mediated peptide thioester ligations, suggesting the use of Fmoc-azidolysine as an orthogonally protected form of lysine, which can be converted to the native amine under reducing conditions.[38] In our hands, incorporation of this residue within peptides yields even more hydrophobic sequences than Fmoc-Lys(Alloc)-OH.

One conceivable solution to this problem is to synthesize a novel protecting group, which retains the positive charge of Lys residues in the form of a non-nucleophilc tertiary amine. Remarkably, only one example of this type of amino protection has been described in the context of peptide synthesis according to our knowledge, and has been used only sparsely in the literature not for the purpose of increasing peptide polarity.[39] **Figure 9** shows the structure of Fmoc-Lys(Pydec)-OH (**52**), which bears a carbamate protecting group conjugated to a 2-pyridyl disulfide moiety. This protecting group can be conveniently removed upon treatment with dithiothreitol or TCEP in aqueous buffer, or upon the action of ethanethiol and DBU in organic solvent. In our hands, incorporation of **52** does seem to afford sequences with improved polarity (marked by faster elution on reverse-phase HPLC). However, preliminary experiments revealed that Lys(Pydec) is unstable within thioacid-containing peptides. This is likely due to the highly activated nature of the pyridyl disulfide bond. It would also be preferable to use a more basic tertiary amine to maintain cationic charge. A few hypothetical examples of such charge-retaining protecting groups for lysine or N-terminal modification are shown in **Figure 10**, each bearing a tertiary amine and different linkers. Potential methods for cleavage are highlighted, including redox-mediated deprotection of **53**, organometallic deprotection of **54**, or cleavage of **55** upon exposure to stronger acids such as trifluoromethanesulfonic acid (TfOH). Potential

application of this strategy to peptide synthesis is highlighted in **Figure 10**. The protecting group can be incorporated directly by SPPS using the corresponding protected lysine residue. Alternatively, the *para*-nitrophenylcarbamate (**60**) could be prepared and used to directly protect lysine residues within a larger expressed protein, while maintaining overall charge. This would enable subsequent isonitrile-mediated ligation at the C-terminus of an expressed protein without reliance on cysteine for expressed-protein ligation.[40]

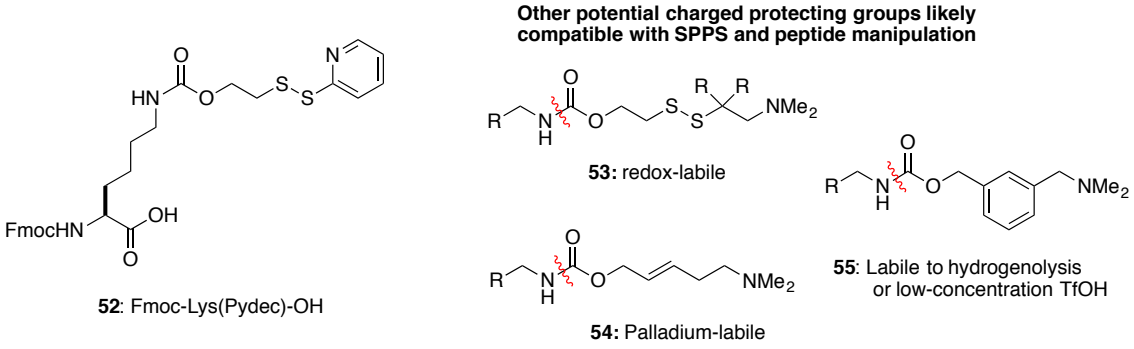


Figure 9: Structure of Lys(Pydec)-OH (**52**), and proposed Lys protecting groups which retain cationic charge.

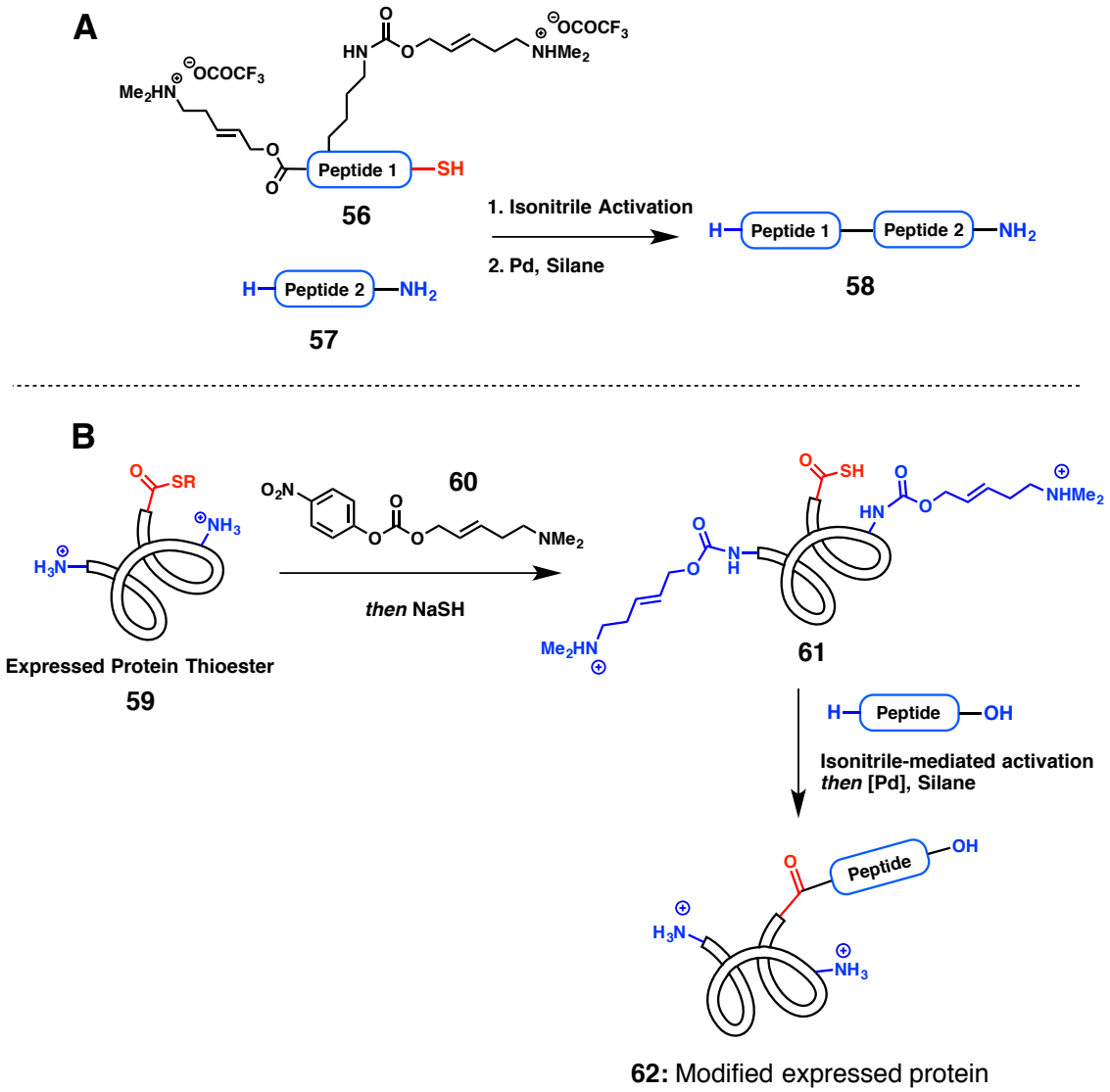
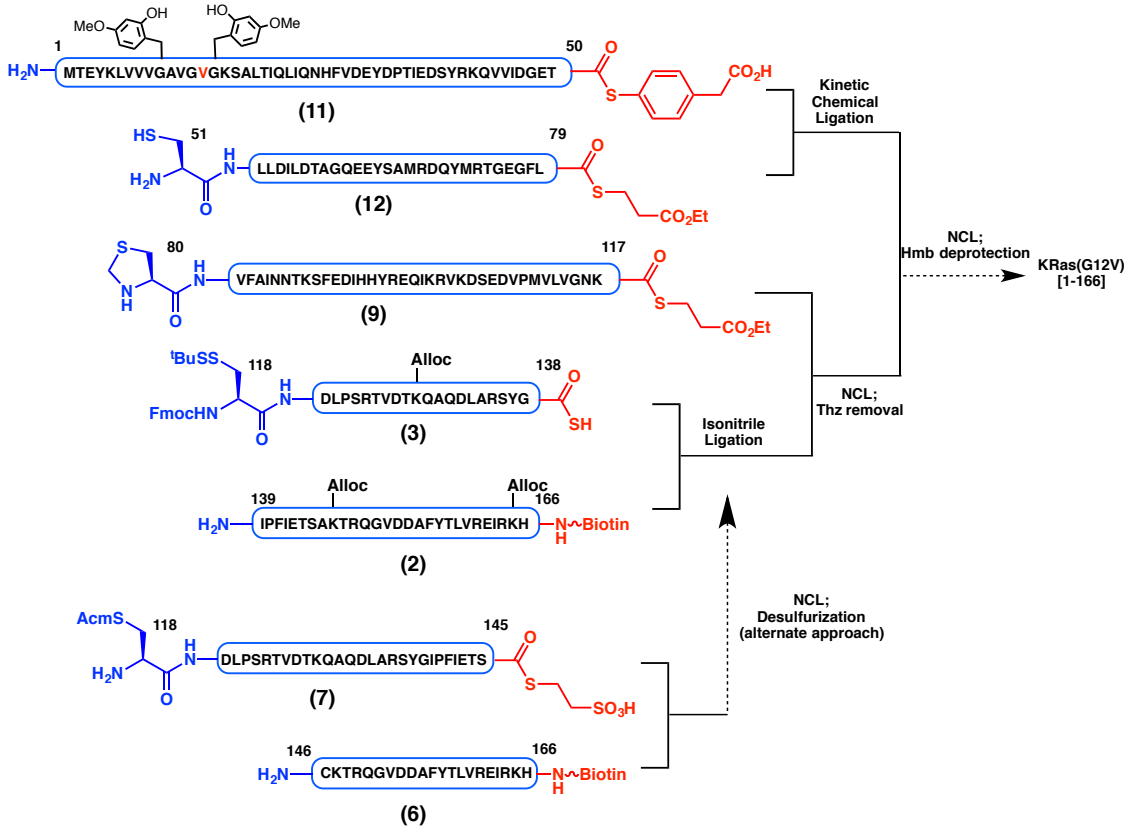


Figure 10: Potential application of charge-retaining Lys protecting groups. (**A**) Application to isonitrile-mediated ligation; (**B**) Direct protection of expressed protein side-chains, and reaction of C-terminal thioacid with peptide nucleophile to produce modified expressed proteins without reliance on cysteine.

EXPERIMENTAL SECTION

Table of Contents:

Part I. Outline of Synthetic Scheme Towards Biotinylated KRas(G12V) [1-166]



Materials and General Procedures.

General Considerations. All commercially available materials were used without purification (Aldrich, Novabiochem, TCI America). 2,2'-Azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride (VA-044) was purchased from Wako Pure Chemical Industries. HATU was purchased from Genscript (Piscataway, New Jersey). Bond-Breaker™ (0.5 M aqueous TCEP, pH = 7.0) solution was purchased from ThermoScientific. L-Ribose was purchased from Carbosynth Limited. All solvents were reagent grade or HPLC grade. All reactions were performed under an atmosphere of purified dry Argon (Ar) unless otherwise indicated. Low-resolution mass spectral analyses were performed with a JOEL JMS_DX_303-HF mass spectrometer or Waters Micromass ZQ mass spectrometer. Analytical TLC was performed on E. Merck silica gel 60 F254 plates and flash column chromatography was performed on E. Merck silica gel 60 (40-63 μm). Yields refer to chromatographically pure compounds.

UPLC-MS, HPLC Methods for Analysis and Purification:

UPLC-MS analysis. UPLC-MS analyses were performed using a WatersAcquity™ Ultra Performance LC systed equipped with Acquity UPLC® BEH C18, 1.7 μL, 2.1 x 100 mm, Acquity UPLC® C8, 1.7 μL, 2.1 x 100 mm, Acquity UPLC® C4, 1.7 μL, 2.1 x 100 mm columns at a flow rate of 0.3 mL/min.

Preparative HPLC. All separations involved a mobile phase of 0.05 v/v% TFA in water (solvent A) and 0.05% v/v% TFA in acetonitrile (solvent B). Preparative separations were performed using a Rainin HPLC solvent delivery system equipped with a Rainin UV-1 detector and Agilent Dynamax reverse phase HPLC column Microsorb 100-8 C18 (250 x 21.4 mm), or Microsorb 300-5 C8 (250 x 21.4 mm) or Microsorb 300-5 C4 (250 x 21.4 mm). Alternatively, xBridge Prep C8 5μm OBD (19 x 150 mm) or xBridge Prep C4 5μm OBD (19 x 150 mm) columns were employed.

Solid-Phase Peptide Synthesis (SPPS) by Fmoc-Strategy. Automated SPPS was performed on an Applied Biosystems Pioneer continuous flow peptide synthesizer or a Biotage Alstra Microwave Peptide Synthesizer. Peptides were synthesized using

automated Fmoc- protocols. The deblock mixtures consisted of a mixture of 96:2:2 of DMF / DBU / piperidine, or 20% piperidine in DMF with or without 0.1 M Oxyma Pure. The appropriate Boc- or Fmoc- amino acids and Fmoc- protected pseudoproline dipeptides (obtained from Novabiochem® in the case of all-L KRas, or synthesized as described in the text in the case of all-D KRas) were employed as indicated. C-terminal thioesters were prepared with using a modification of elongation methods as described.

Standard Amino Acid Residues Used.

Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Cys(StBu)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Cys(Acm)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(OtBu), Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Met-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Val-OH, Boc-Thz-OH.

Representative Microwave-Assisted SPPS Protocol.

Prior to synthesis: Swell resin (0.1 mmol scale, DMF, 50°C, 10 min)

Deprotection: add 20% Piperidine in DMF (5 mL, 50°C, 5 min) / Wash with DMF.

> *Or* 20% Piperidine in DMF containing 0.1M OxymaPure (5 mL, 3 min RT, then replace solution, 5 min RT) / Wash with DMF

Coupling: add amino acid (4 equiv, 0.5 M in DMF), add OxymaPure (4 equiv, 0.5 M in DMF), add DIC (4 equiv, 0.5 M in DMF), additional DMF added (to reach 3 mL volume), coupling (room temp, 2 min, then 50°C for 8 minutes) / Wash with DMF

Peptide Cleavage and Deprotection:

Mild Cleavage from Trityl Resin to Access Fully Protected Peptides. Upon completion of automated synthesis on a 0.1 mmol scale, the peptide resin was washed into a peptide synthesis vessel with CH_2Cl_2. The resin was then subjected to a cleavage cocktail (1:1:8 acetic acid / TFE / CH_2Cl_2) for 30 min (3 x 10 mL), filtering after each

treatment. The combined cleavage solutions were then concentrated under reduced pressure to a minimum volume of acetic acid, precipitated with water, shell frozen, and lyophilized.

Acid Labile Protecting Group Removal. Peptides were subjected to cocktail B (~3 mL / 100 mg of peptide) consisting of TFA (88% by volume), water (5 % by volume), phenol (5% by weight), and triisopropylsilane (2% by volume). Alternatively, peptides were subjected to cocktail R, consisting of TFA (90% by volume), Thioanisole (5% by volume), 1,2-ethanedithiol (3% by volume), and anisole (2% by volume). The resulting solution was either concentrated to half its original volume with Ar stream and precipitated with ice-cold diethyl ether (10 volume equivalents) or directly precipitated upon the addition of ice-cold diethyl ether (10 volume equivalents). Centrifugation of the resulting precipitate and decantation afforded a white pellet, which was further washed twice with ice-cold diethyl ether, centrifuged, and decanted as before. The solid was then dissolved in water/acetonitrile (1:1, 0.1% TFA) and lyophilized. The resulting solid was purified by RP-HPLC.

Kinetic Chemical Ligation (KCL) Buffer. The buffer required for kinetic chemical ligation (KCL) was freshly prepared prior to the reaction. Guanidine·HCl (1.146 g, 12 mmol), Na_2HPO_4 (56.6 mg, 0.4 mmol), and TCEP·HCl (10.8 mg, 0.04 mmol) were weighed into a 14 mL centrifuge vial. The resulting solids were solubilized in H_2O (2 mL), and the pH was adjusted to the desired pH with 5M NaOH. The resulting solution was degassed by argon sparge with sonication for at least 15 minutes before use.

Native Chemical Ligation Buffer.

The buffer required for native chemical ligation (NCL) was freshly prepared prior to the reaction. Guanidine·HCl (1.146 g, 12 mmol), Na_2HPO_4 (56.6 mg, 0.4 mmol), and TCEP·HCl (10.8 mg, 0.04 mmol), and 4-mercaptophenylacetic acid (MPAA) (67 mg, 0.4 mmol) were weighed into a 14 mL centrifuge vial. The resulting solids were solubilized in H_2O (1.8 mL), the pH was adjusted to the desired pH (generally 6.8-7.2)

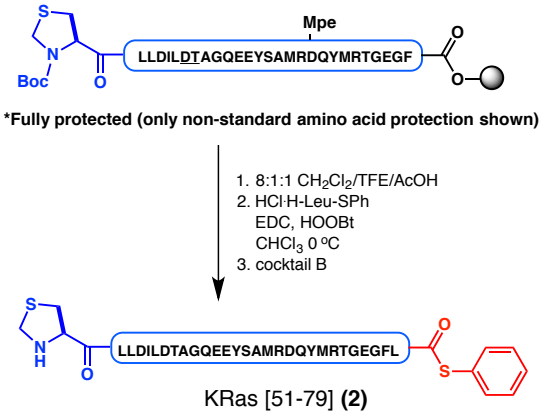
with 5M NaOH, and the final volume adjusted to reach 2 mL. The resulting solution was degassed by argon sparge with sonication for at least 15 minutes before use.

Representative procedure for first-residue loading of NovaSyn® TGT alcohol resins:

A suspension of Novasyn TGT Trityl-OH resin (1.36 g, 0.258 mmol, 0.19 mmol/g loading) in anhydrous toluene (12 mL) was treated with AcCl (1.40 mL) and heated at 60°C for 3.5 hours with moderate agitation. The resin was then filtered into an oven-dried peptide vessel, and washed with anhydrous toluene (2 x 15 mL) followed by CH_2Cl_2 (3 x 15 mL), and finally the vessel was capped with a rubber septum and placed under an argon atmosphere. Separately, Fmoc-D-Glu(OtBu)-OH (276 mg, 0.65 mmol) was dissolved in anhydrous CH_2Cl_2 (12 mL) and *N,N*-diisopropylethylamine (0.45 mL, 2.58 mmol). This solution was carefully added to the washed resin via syringe. The reaction vessel was agitated for 2.5 hours at room temperature under argon. The solvent was then filtered with nitrogen pressure, and the resin was washed with 17:2:1 CH_2Cl_2/MeOH/DIPEA (2 x 20 mL, 10 min each). The resin was further washed with CH_2Cl_2 (3 x 15 mL), then DMF (3 x 15 mL), then finally CH_2Cl_2 (2 x 15 mL) and dried under high vacuum. Fmoc loading was estimated by the method of Gude and co-workers[41] following cleavage with 2% DBU in DMF and measurement of UV absorbance at 304 nm, and found to be 0.16 mmol/g.

Part III: Synthesis of Peptidyl Fragments, Thioesters and Thioacids

Synthesis of KRas[51-79] aryl thioester 2



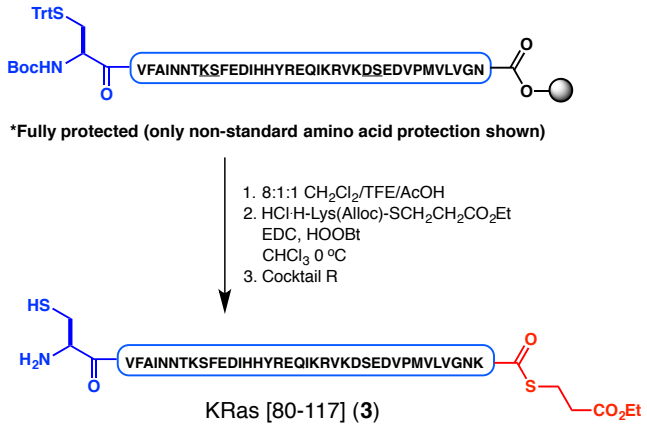
Synthesized on a 0.1 mmol scale starting from Fmoc-Phe-OH loaded NovaSyn TGT resin. Non-standard amino acids (used at the indicated positions) were Fmoc-Asp(^{t}Bu)-Thr($\Psi^{Me,Me}$pro)-OH, and Fmoc-Asp(Mpe)-OH. Cleavage from resin using 8:1:1 CH_2Cl_2/TFE/AcOH according to the general protocol afforded the crude fully protected peptide.

To a 0 ^{o}C solution of crude protected peptide (384 mg, 69.1 μmol) in $CHCl_3$ (2.5 mL) was added HCl-H-Leu-SPh[42] (51 mg, 0.196 mmol) followed by HOOBt (0.319 mmol) and EDC (37 μL, 0.209 mmol). The reaction was stirred in an ice bath for 2 hours, at which point UPLC/MS analysis indicated a small amount of residual starting material. Additional EDC (30 μL, 0.169 mmol) was added and the reaction was stirred for a further 40 minutes. At this point, the reaction was quenched upon the addition of 8 mL $CHCl_3$ (+5% AcOH). The organic mixture was washed with H_2O (3 mL), and then concentrated. Global deprotection with cocktail B (8 mL, 3.75 hours, argon) followed by precipitation with diethyl ether according to the general procedure afford the crude peptide. The solid residue was further purified by RP-HPLC to afford KRas[51-79] (**2**) as a white solid (40.0 mg, 17% yield).

HPLC conditions: C8 X-bridge column 5 μm OBD, 20-70% MeCN (+0.05% TFA) over 30 min. *Flow rate:* 16 mL/min. T_R = 16 min.

ESI-MS(+) for peptide **2**. Chemical Formula = $C_{149}H_{225}N_{37}O_{48}S_4$. MW = 3430.89 g/mol. ESI calculated for $[M+2H^+]^{2+}$ *m/z*: 1716.44, found: 1716.02; $[M+3H^+]^{3+}$ *m/z*: 1144.63, found: 1144.46;

Synthesis of KRas[80-117] (3)



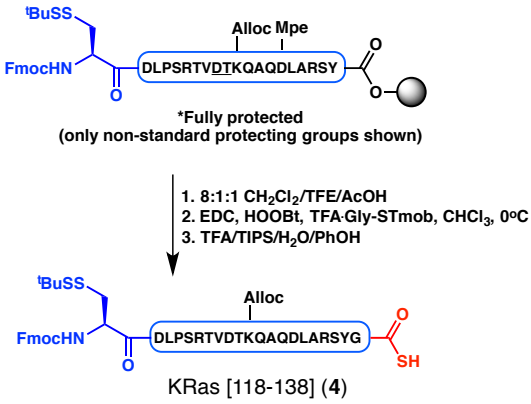
Synthesized on a 0.1 mmol scale starting from Fmoc-Asn(Trt)-OH loaded NovaSyn TGT resin. Non-standard amino acids (used at the indicated positions) were Fmoc-Lys(Boc)-Ser($\Psi^{Me,Me}$pro)-OH and Fmoc-Asp(^{t}Bu)-Ser($\Psi^{Me,Me}$pro)-OH. Cleavage from resin using 8:1:1 CH_2Cl_2/TFE/AcOH according to the general protocol afforded the crude fully protected peptide.

To a 0 ^{o}C solution of the crude protected peptide (215 mg, 28.8 µmol) in $CHCl_3$ (1.3 mL) was added HCl·H-Lys(Alloc)-SAlkyl **S-19** (31 mg, 68.7 µmol) and HOOBt (19 mg, 0.117 mmol) followed by EDC (15 µL, 84.8 µmol). The reaction was stirred for 1 hour, at which point UPLC/MS analysis indicated full conversion. The reaction was warmed to room temperature, and $PhSiH_3$ (88 µL, 0.71 mmol) was added followed by $Pd(PPh_3)_4$ (12 mg, 10.4 µmol). The reaction was stirred for 1 hour, and was quenched upon the addition of 6 mL $CHCl_3$ (+5% AcOH). The organic mixture was washed with H_2O (2 mL) and then concentrated. Global deprotection using Cocktail R (6 mL, 2.5 hr, argon) and precipitation according to the general procedure yielded the crude peptide. Purification by RP-HPLC afforded **3** as a white solid (42 mg, 32% yield).

HPLC conditions: C8 xBridge column 5 µm OBD, 10-60% MeCN (+0.05% TFA) over 30 min. *Flow rate*: 16 mL/min.

ESI-MS(+) for peptide **3**. Chemical Formula = $C_{201}H_{320}N_{56}O_{60}S_3$. MW = 4577.28 g/mol. ESI calculated for $[2M+5H^+]^{5+}$ *m/z*: 1831.91, found: 1831.57; $[M+3H^+]^{3+}$ *m/z*: 1526.76, found: 1526.54; $[M+4H^+]^{4+}$ *m/z*: 1145.53, found: 1145.05; $[M+5H^+]^{5+}$ *m/z*: 916.45, found: 916.30; $[M+6H^+]^{6+}$ *m/z*: 763.88, found: 763.83; $[M+7H^+]^{7+}$ *m/z*: 654.89, found: 654.78.

Synthesis of KRas(G12V) [118-138] thioacid (4):



Synthesized on a 0.1 mmol scale using Fmoc-Tyr(tBu)-OH Novasyn TGT resin. Non-standard amino acids (used at the indicated positions) were Fmoc-Asp(OMpe)-OH, Fmoc-Asp(tBu)-Thr($\Psi^{Me,Me}$pro)-OH, and Fmoc-Lys(Alloc)-OH. Following synthesis, the resin was cleaved under mild conditions using 8:1:1 CH_2Cl_2:TFE:AcOH according to the general procedure to afford the crude protected peptide.

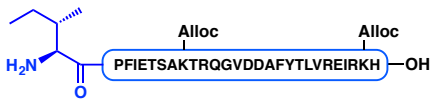
To a 0°C solution of this protected peptide (269 mg, 65.4 µmol) in $CHCl_3$ (2.50 mL) was added HOOBt (49.0 mg, 0.30 mmol) followed by TFA-Gly-STmob[7] (74.0 mg, 0.19 mmol) and EDC (53 uL, 0.30 mmol). After stirring at this temperature for 1 hour, the reaction was quenched by the addition of 5% AcOH in $CHCl_3$ (10 mL). The organic layer was washed with H_2O (4 mL), and the aqueous layer was extracted once more with 5 mL 5% AcOH solution in $CHCl_3$. The combined organic layers were concentrated and subjected to cocktail B deprotection (9 mL, 2.5 hours, argon). The

peptide was directly precipitated by the addition of cold ether, the mixture was centrifuged and ether was decanted. The ether wash was repeated once more, and the resulting residue was dissolved in 1:1 MeCN/H_2O and lyophilized. RP-HPLC purification of the residue afforded thioacid **4** as a white solid (45.0 mg, 26% yield).

HPLC conditions: xBridge Prep C8 column 5µm OBD (19 x 150 mm). 30-70% MeCN/H_2O (0.05% TFA) over 30 min. *Flow Rate:* 15 mL/min. T_r = 13.3 min.

ESI-MS(+) for peptide **4**: Chemical Formula = $C_{119}H_{180}N_{30}O_{38}S_3$. MW = 2735.10 g/mol. ESI calculated for $[2M+3H^+]^{3+}$ *m/z:* 1824.40, found 1824.00; $[M+2H^+]^{2+}$ *m/z:* 1368.55, found: 1368.32; $[M+3H^+]^{3+}$ *m/z:* 912.70, found: 912.35.

Synthesis of KRas[139-166] (5)



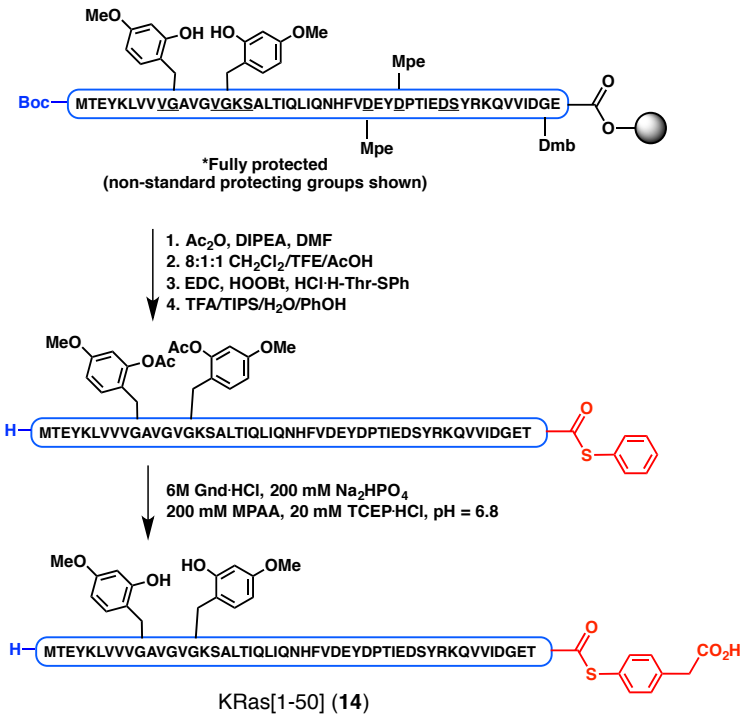
KRas [139-166] **(5)**

Synthesized on a 0.2 mmol scale, starting with Fmoc-His(Trt)-OH NovaSyn TGT Resin. Resin cleavage and global deprotection was accomplished using Cocktail R (20 mL, 2 hours). The peptide was precipitated with cold diethyl ether according to the general procedure. Purification of a portion of this crude material by RP-HPLC afforded **5** as a white solid (21.5 mg)

HPLC conditions: C8 xBridge column 5 µm OBD, 15-60% MeCN (+0.05% TFA) over 30 min. *Flow rate*: 16 mL/min. T_R = 18.0 min.

ESI-MS(+) for peptide **5**. Chemical Formula = $C_{156}H_{244}N_{42}O_{47}$. MW = 3459.92 g/mol. ESI calculated for $[M+2H^+]^{2+}$ *m/z*: 1730.96, found: 1730.48; $[2M+5H^+]^{5+}$ *m/z*: 1384.96, found: 1384.55; $[M+3H^+]^{3+}$ *m/z*: 1154.30, found: 1153.83; $[M+4H^+]^{4+}$ *m/z*: 865.98, found: 865.75; $[M+5H^+]^{5+}$ *m/z*: 692.98, found: 692.73.

Synthesis of KRas(G12V) [1-50] MPAA thioester (14):



Synthesized on a 0.1 mmol scale using a Biotage Alstra microwave peptide synthesizer, using Fmoc-Glu(^{t}Bu)-OH Novasyn TGT resin. Non-standard amino acids (used at the indicated positions) were Fmoc-(Dmb)Gly-OH, Fmoc-Val-(Hmb)Gly-OH dipeptide, Fmoc-Asp(Mpe)-OH, Fmoc-Lys(Boc)-Ser($\Psi^{Me,Me}$pro)-OH and Fmoc-Asp(O^{t}Bu)-Ser($\Psi^{Me,Me}$pro)-OH. The peptide was synthesized according to the general microwave SPPS protocol using 20% Piperidine in DMF as deblock, with the following variations: (a) The residue after Fmoc-(Dmb)Gly-OH was double coupled, each cycle for 30 minutes. (b) Coupling of Fmoc-Val-(Hmb)Gly-OH dipeptide was performed for 45 minutes.

After synthesis, the Hmb groups were acetylated using the following procedure: The resin (assumed 0.1 mmol) was swelled in 10 mL DMF for 10 minutes, and the solution drained. To the resin was added a premade solution of Ac_2O (0.19 mL, 2.0 mmol) and DIEA (0.4 mL, 2.3 mmol) in DMF (8.0 mL). After 2 hours, the solution was drained, and the resin was washed with 3 x 10 mL DMF. Following acetylation, the resin was

cleaved under mild conditions using 8:1:1 CH_2Cl_2:TFE:AcOH according to the general procedure to afford the crude protected peptide.

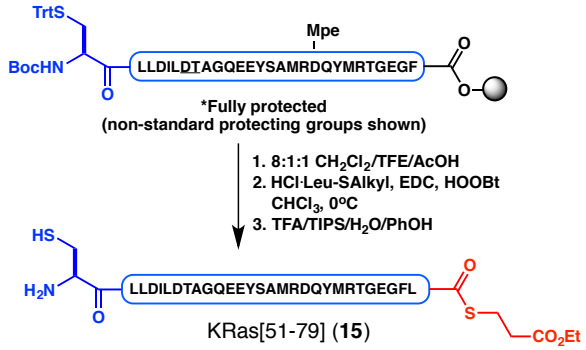
To a solution of this protected peptide (260 mg, 29.8 μmol) in $CHCl_3$ (1.5 mL) was added HOOBt (19.4 mg, 0.119 mmol) and HCl-D-Thr-SPh[43] (19.2 mg, 0.078 mmol). The mixture was cooled to 0°C, and EDC (21.0 uL, 0.119 mmol) was added. The reaction was stirred for 1.5 hours, at which point additional EDC (10 uL, 0.056 mmol) was added. After 45 minutes, the reaction was quenched by the addition of 5% AcOH in $CHCl_3$. The mixture was washed with H_2O (4 mL), and the aqueous phase was extracted once with 5% AcOH/$CHCl_3$ (5 mL). The combined organic layers were concentrated, and subjected to cocktail B deprotection (8 mL, 2.5 hr, argon balloon). The peptide was precipitated with cold ether (~60 mL), and the mixture centrifuged. The resulting pellet was dissolved in 1:1 MeCN/H_2O (0.1 % TFA) and lyophilized to afford the crude peptide (189 mg).

A portion of the obtained crude peptide (88 mg) was dissolved in 10.0 mL argon-degassed thiol buffer (200 mM MPAA, 200 mM Na_2HPO_4, 6M Gn·HCl, 20 mM TCEP·HCl, pH~6.8). The mixture was stirred for 8 hours at room temperature under argon, at which point the acetyl groups had been fully deprotected by UPLC/MS analysis. The reaction was directly subjected to RP-HPLC purification to afford **14** as a white solid (15.4 mg, 19% yield overall).

HPLC conditions: Microsorb 300-5 C4 (250 x 21.4 mm) column. 30-65% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow rate:* 20 mL/min. T_r=12.5 min.

ESI-MS(+) of peptide **14**: Chemical Formula = $C_{273}H_{418}N_{62}O_{85}S_2$. MW = 5992.82 g/mol. ESI calcd. for $[M+3H^+]^{3+}$ *m/z*: 1998.61, found: 1998.83; $[2M+7H^+]^{7+}$ *m/z:* 1713.23, found: 1713.50; $[M+4H^+]^{4+}$ *m/z:* 1499.20, found: 1499.08; $[M+5H^+]^{5+}$ *m/z:* 1199.59, found: 1199.46, $[M+6H^+]^{6+}$ *m/z:* 999.80, found: 999.91; $[M+7H^+]^{7+}$ *m/z:* 857.12, found: 857.42.

Synthesis of KRas(G12V) [51-79] Alkyl Thioester (15):



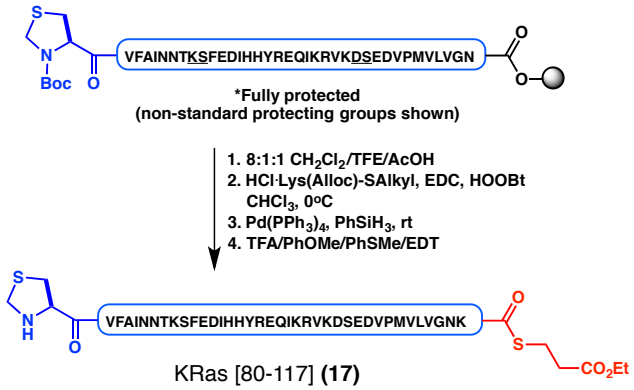
Synthesized on a 0.1 mmol scale using Fmoc-Phe-OH Novasyn TGT resin. Non-standard amino acids (used at the indicated positions) were Fmoc-Asp(Mpe)-OH and Fmoc-Asp(tBu)-Thr($\Psi^{Me,Me}$pro)-OH. Following synthesis, the resin was cleaved under mild conditions using 8:1:1 CH_2Cl_2:TFE:AcOH according to the general procedure to afford the crude protected peptide.

To a 0°C solution of this protected peptide (333 mg, 64.4 μmol) in $CHCl_3$ (3.5 mL) was added HCl-H-D-Leu-SAlkyl **S-18** (55.2 mg, 0.195 mmol) and HOOBt (37.5 mg, 0.23 mmol), followed by the addition of EDC (40.6 μL, 0.23 mmol). The reaction was stirred at this temperature for 90 minutes, and was quenched upon the addition of 5% AcOH in $CHCl_3$ (~6 mL). The reaction mixture was washed with H_2O (3 mL), the aqueous layer was extracted once with 5% AcOH in $CHCl_3$ (5 mL), and the combined organic layers were directly concentrated. The residue was subjected to cocktail R deprotection (8 mL, 2 hours, argon balloon). The peptide was directly precipitated upon the addition of cold diethyl ether (80 mL) and the solids were collected upon centrifugation. The resulting pellet was further washed with cold diethyl ether (2 x 20 mL), suspended in 1:1 MeCN/H_2O (0.1% TFA), and lyophilized. Purification of the residue by RP-HPLC afforded **15** as a white solid (50.0 mg, 22% yield).

HPLC Conditions: Microsorb 300-5 C8 (250 x 21.4 mm) column. 25-75% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow Rate*: 20 mL/min. T_r = 15.5 minutes.

ESI-MS(+) for peptide **15**: Chemical Formula = $C_{147}H_{229}N_{37}O_{50}S_4$. MW = 3442.90 g/mol. ESI calculated for $[M+2H^+]^{2+}$ *m/z:* 1722.45, found: 1722.35; $[M+3H^+]^{3+}$ *m/z:* 1148.63, found: 1148.55; $[M+4H^+]^{4+}$ *m/z*: 861.72, found: 861.67.

Synthesis of KRas(G12V) [80-117] Alkyl Thioester (17):



Synthesized on a 0.1 mmol scale starting with Fmoc-Asn(Trt)-OH Novasyn TGT resin. Non-standard amino acids (used at the indicated positions) were Fmoc-Asp(tBu)-Ser($\Psi^{Me,Me}$pro)-OH, and Fmoc-Lys(Boc)-Ser($\Psi^{Me,Me}$pro)-OH. Following synthesis, the resin was cleaved under mild conditions using 8:1:1 CH_2Cl_2:TFE:AcOH according to the general procedure to afford the crude protected peptide.

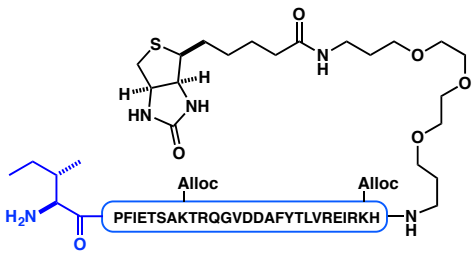
To a 0°C solution of this crude protected peptide (all L-amino acids, 248 mg, 0.034 mmol) in $CHCl_3$ (1.6 mL) was added HOOBt (21.2 mg, 0.13 mmol) and HCl-Lys(Alloc)-SAlkyl (**S-19**) (35.0 mg, 0.091 mmol). EDC (17.0 µL, 0.096 mmol) was then added, and the reaction was stirred at this temperature for 40 minutes, at which point the reaction was complete by UPLC/MS analysis. The mixture was warmed to room temperature for 15 minutes, and $PhSiH_3$ (99 µL, 0.80 mmol) and $Pd(PPh_3)_4$ (14.0 mg, 0.012 mmol) were added. The reaction was stirred for 2 hours, and was then diluted with 5% AcOH/$CHCl_3$ (8 mL). The mixture was washed with H_2O (2 mL) and extracted once further with 5% AcOH/$CHCl_3$ (~5 mL). The combined organic layers were directly concentrated and subjected to cocktail R deprotection (6.5 mL, 1.5 hours, argon balloon). The peptide was precipitated by the addition of cold ether, the mixture centrifuged, and ether was decanted. The resulting pellet was dissolved in 1:1

MeCN/H_2O (0.1% TFA), shell-frozen and lyophilized. The crude peptide was purified by HPLC to obtain **17** (35.0 mg, 22% yield) as a white solid.

HPLC Conditions: xBridge Prep C8 column 5µm OBD (19 x 150 mm), 20-60% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow Rate*: 15 mL/min. T_r = 8.2 minutes.

ESI-MS(+) for compound **17**: Chemical Formula = $C_{202}H_{320}N_{56}O_{60}S_3$. MW = 4589.29 g/mol. ESI calculated for $[2M+5H^+]^{5+}$ *m/z*: 1836.71, found: 1836.15; $[M+3H^+]^{3+}$ *m/z:* 1530.76, found: 1530.29; $[M+4H^+]^{4+}$ *m/z*: 1148.32, found: 1147.83; $[M+5H^+]^{5+}$ *m/z*: 918.86, found: 918.62; $[M+6H^+]^{6+}$ *m/z*: 765.88, found: 765.63

Synthesis of KRas(G12V) [139-166] (22) with biotin label:



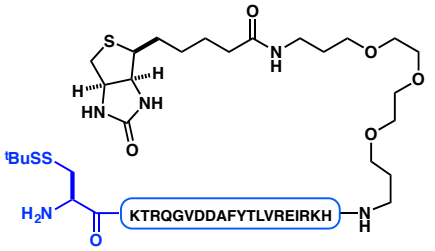
KRas [139-166] **(19)**

Synthesized on 0.05 mmol scale beginning with Fmoc-PEG Biotin NovaTag™ resin (purchased from Novabiochem). The resin was directly cleaved by the addition of Cocktail R (90/5/3/2 TFA/thioanisole/1,2-ethanedithiol/anisole) for 2.5 hours under argon and precipitated with ether according to the general protocol. The residue was purified by RP-HPLC to afford **19** as a white solid (20.7 mg, 11% yield).

HPLC conditions: xBridge Prep C8 column 5µm OBD (19 x 150 mm), 20-70% MeCN (0.05% TFA) against H_2O (0.05% TFA) over 30 min. *Flow Rate:* 15 mL/min. T_r = 13.8 min.

ESI-MS(+) for peptide **19**: Chemical Formula = $C_{176}H_{280}N_{46}O_{51}S$. MW = 3888.51 g/mol. ESI calculated for $[M+2H^+]^{2+}$ *m/z:* 1945.25, found: 1944.92; $[2M+5H^+]^{5+}$ *m/z*: 1556.40, found: 1556.06; $[M+3H^+]^{3+}$ *m/z:* 1297.17, found: 1296.87; $[M+4H^+]^{4+}$ *m/z*: 973.12, found: 972.92; $[M+5H^+]^{5+}$ *m/z*: 778.70, found: 778.62.

Synthesis of KRas[146-166] with biotin label (22):



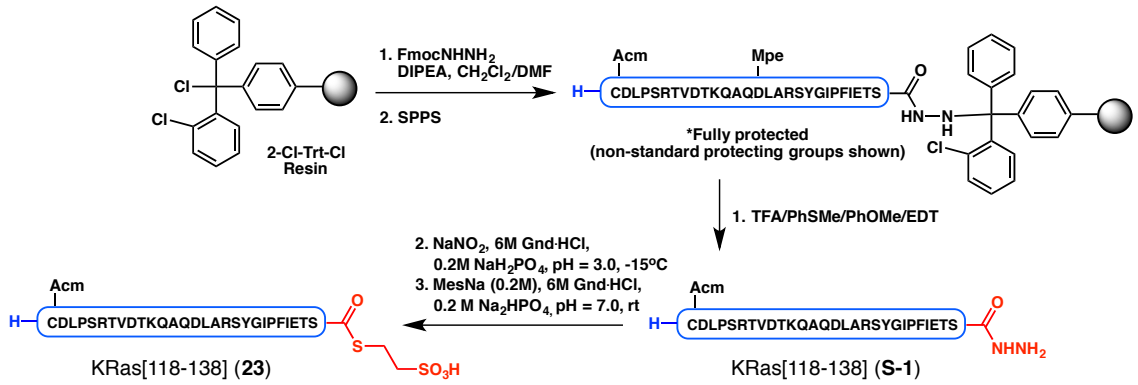
KRas [146-166] **(22)**

Synthesized on 0.1 mmol scale beginning with Fmoc-PEG Biotin NovaTag™ resin (purchased from Novabiochem). This fragment was synthesized using the microwave SPPS general procedure. Deblock used was 20% piperidine in DMF plus 0.1M oxymapure. The resin was directly cleaved by the addition of Cocktail R (90/5/3/2 TFA/thioanisole/1,2-ethanedithiol/anisole) for 1.5 hours under argon and precipitated with ether according to the general protocol. The residue was purified by RP-HPLC to afford **22** as a white solid (42.0 mg, 14% yield).

HPLC conditions: xBridge Prep C8 column 5µm OBD (19 x 150 mm). 10-65% MeCN/H_2O (0.05% TFA) over 30 min. *Flow Rate:* 16 mL/min. T_r = 15.4 min.

ESI-MS(+) for peptide **22**: Chemical Formula = $C_{134}H_{223}N_{39}O_{36}S_3$. MW = 3052.68 g/mol. ESI calculated for $[M+2H^+]^{2+}$ *m/z:* 1527.34, found: 1526.90; $[2M+5H^+]^{5+}$ *m/z*: 1222.07, found: 1221.83; $[M+3H^+]^{3+}$ *m/z:* 1018.56, found: 1018.23; $[M+4H^+]^{4+}$ *m/z*: 764.17, found: 764.21; $[M+5H^+]^{5+}$ *m/z*: 611.53, found: 611.35.

Synthesis of KRas(G12V) [118-138] Thioester (23):



23 was synthesized using the hydrazide oxidation method.[44] 2-Cl-Trt-NHNHFmoc resin was synthesized from 2-Cl-Trt-Cl resin using a modified procedure, initially described by Huang, et al.[45] To a 30 mL oven-dried peptide vessel was added 2-Cl-Trt-Cl resin (1.0 g, 1.5 mmol, loading ~ 1.5 mmol/g). The resin was swelled in 10 mL anhydrous CH_2Cl_2 for 20 minutes under argon. At this time, the solvent was drained under N_2 pressure, and the resin was resuspended in anhydrous CH_2Cl_2 (8 mL). Separately, 9-fluorenylmethyl carbazate (127.5 mg, 0.5 mmol) and *N,N*-diisopropylethylamine (0.34 mL, 2.0 mmol) were dissolved in anhydrous DMF (8 mL). This solution was slowly added to the suspended resin, and the peptide vessel was gently agitated at room temperature overnight under argon. The solution was filtered under N_2 pressure, and the resin was washed with 17:2:1 CH_2Cl_2/MeOH/DIPEA (2 x 10 mL, for 10 minutes each). The resin was then consecutively washed with CH_2Cl_2 (2 x 8 mL), DMF (2 x 8 mL), CH_2Cl_2 (3 x 8 mL), MeOH (1 x 8 mL), and finally diethyl ether (2 x 8 mL), and the resin was dried under high vacuum. Fmoc loading was spectroscopically determined to be approximately 0.40 mmol/g.

The corresponding acyl hydrazide fragment **S-1** was synthesized on a 0.1 mmol scale using the general procedure for microwave-assisted SPPS. Deblock – 20% piperidine in DMF (containing 0.1M Oxyma), room temperature. Non-standard amino acids (used at the indicated positions) were Fmoc-Asp(Mpe)-OH and Fmoc-Cys(Acm)-OH. Cleavage from resin was accomplished using cocktail R (8 mL, 2 hours) followed by precipitation

with ether and centrifugation according to the general protocol. The crude peptide was purified by RP-HPLC. Purest fractions were combined, shell-frozen, and lyophilized to afford peptide hydrazide **S-1** (88.3 mg, 27% yield) as a white solid.

HPLC Conditions*:* xBridge Prep C8 column 5µm OBD (19 x 150 mm), 10-60% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow Rate:* 16 mL/min. T_r = 10.5 minutes.

ESI-MS(+) for peptide **S-1:** Chemical Formula = $C_{137}H_{222}N_{40}O_{46}S$. MW = 3197.58 g/mol. ESI calculated for $[M+2H^+]^{2+}$ *m/z*: 1599.79, found: 1599.47; $[M+3H^+]^{3+}$ *m/z*: 1066.86, found: 1066.49; $[M+4H^+]^{4+}$ *m/z*: 800.39, found: 800.38

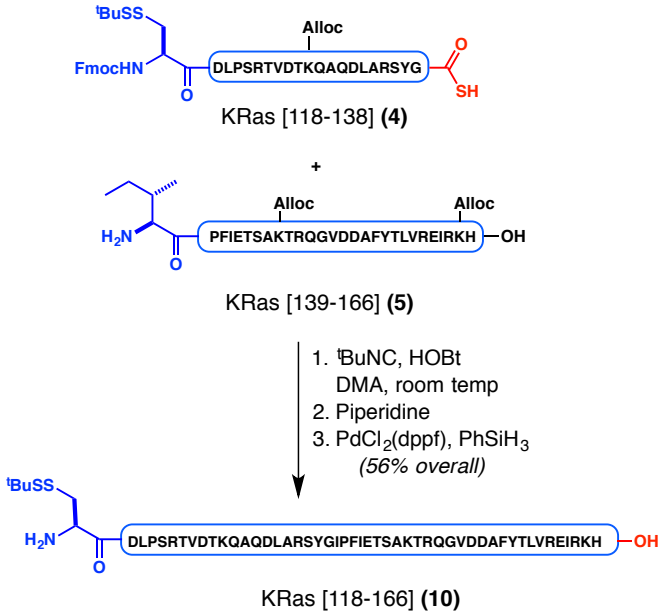
Hydrazide **S-1** (15.2 mg, 4.75 µmol) was dissolved in 950 µL oxidation buffer (6 M Gnd·HCl, 0.2 M NaH_2PO_4, pH=3) in a 15 mL conical plastic centrifuge vial. The vial was briefly centrifuged, and the solution was cooled in a -15°C cooling bath (dry ice/acetone). To this rapidly stirred solution was added $NaNO_2$ (0.5 M in H_2O, 94.8 uL, 47.4 µmol). The reaction was stirred at this temperature for 20 minutes, and was quenched with MesNa (Prepared as a 0.2M solution, containing 6M Gnd·HCl, 0.2 M Na_2HPO_4, pH=7.0, 2.30 mL, 0.46 mmol). The reaction was then removed from the cooling bath and stirred at room temperature for 45 minutes, and purified by RP-HPLC. Purest fractions were combined, shell-frozen, and lyophilized to afford **23** as a white solid (14.5 mg, 92% yield).

HPLC conditions: xBridge Prep C8 column 5µm OBD (19 x 150 mm), 10-60% MeCN in H_2O (+0.05% TFA) over 30 minutes. *Flow Rate*: 16 mL/min. T_r = 13.5 minutes

ESI-MS(+) for peptide **23**: Chemical Formula = $C_{139}H_{224}N_{38}O_{49}S_3$. MW = 3307.72 g/mol. ESI calculated for $[M+2H^+]^{2+}$ *m/z:* 1654.86*,* found: 1654.43; $[M+3H^+]^{3+}$ *m/z*: 1103.57, found: 1103.33.

Part IV: Ligation procedures for the construction of KRas(G12V)

Synthesis of KRas[118-166] (10)



Thioacid 4 (13.4 mg, 4.9 µmol), **4** (21.5 mg, 6.2 mol), and HOBt (8.3 mg, 61 µmol) were weighed directly into a 1-dram glass scintillation vial with a stirbar and screwcap. Anhydrous DMA (1.0 mL) was added, followed by ^{t}BuNC (1.25 mg, 15.0 µmol, dissolved in 400 µL DMA, prepared as stock solution). The reaction was stirred for 72 hours at room temperature, at which point the reaction was transferred via micropipette into a 15 mL conical centrifuge vial. Transfer was quantitated with 100 µL DMA, and piperidine (300 µL) was added. The mixture was stirred for 1 hour, at which point cold diethyl ether (10 mL) was added to precipitate the crude peptides. The vial was centrifuged, ether decanted, and the residue was again washed with ether, and centrifuged as before. The resulting solid was lyophilized from 1:1 MeCN/H_2O (+0.05% TFA).

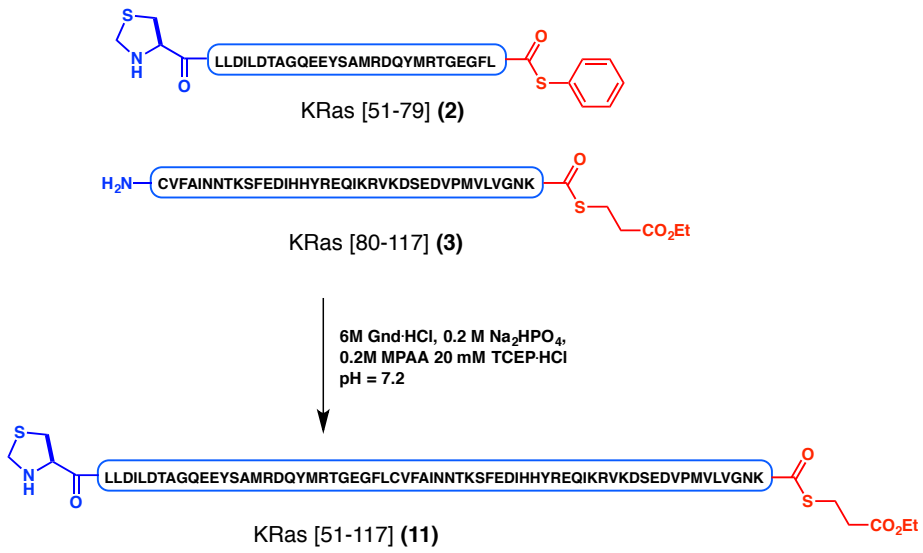
The lyophilized solids were redissolved in DMF (1.50 mL, degassed with argon sparge), and $PhSiH_3$ (29.6 µL, 0.24 mmol) was added followed by $PdCl_2$(dppf) (2.40 mg, 3.0 µmol, dissolved in 100 µL DMF). The reaction was stirred for 15 minutes, at which point UPLC/MS indicated full deprotection. Cold ether (10 mL) was added to precipitate the crude peptides, and the suspension was centrifuged and ether was decanted. The solid residue was dissolved in 1:1 MeCN/H_2O (+0.1% TFA) and purified

by RP-HPLC. Combined fractions were lyophilized to afford **10** as a white solid (16.2 mg, 56% yield overall).

HPLC Conditions: C8 xBridge column 5 µm OBD, 20-70% MeCN (+0.05% TFA) against H_2O (+0.05% TFA) over 30 minutes. Flow rate: 15 mL/min. T_r = 12 min.

ESI-MS(+) for peptide **10**. Chemical Formula = $C_{248}H_{400}N_{72}O_{77}S_2$. MW = 5686.48 g/mol. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1896.49, found: 1895.91; $[2M+7H^+]^{7+}$ *m/z*: 1625.70, found: 1625.32; $[M+4H^+]^{4+}$ *m/z*: 1422.62, found: 1422.20; $[M+5H^+]^{5+}$ *m/z*: 1138.29, found: 1137.85; $[M+6H^+]^{6+}$ *m/z*: 948.74, found: 948.40; $[M+7H^+]^{7+}$ *m/z*: 813.35, found: 813.25; $[M+8H^+]^{8+}$ *m/z*: 711.81, found: 711.63; $[M+9H^+]^{9+}$ *m/z*: 632.83, found: 632.73.

NCL between 2 and 3 to produce KRas[51-117] (11)

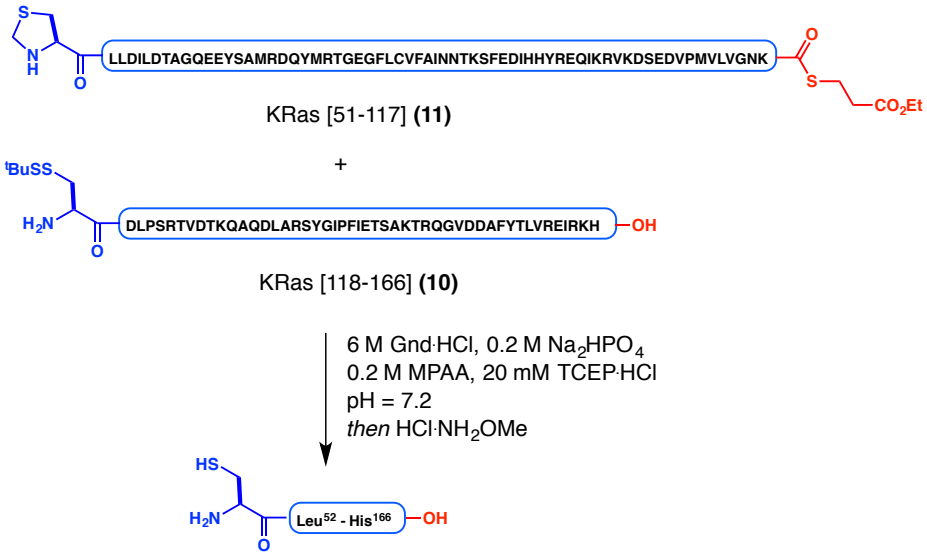


KRas[51-79] **2** (22.0 mg, 6.41 µmol) and KRas [80-117] **3** (41.5 mg, 9.0 µmol) were weighed into an 8 mL glass vial with a stirbar. The peptides were solubilized in ligation buffer (2.60 mL) and stirred under argon for 5 hours. At this point, TCEP (0.5 M, pH = 7.0, 30 µL, 15 µmol) was added and the reaction was stirred for 10 minutes. The mixture was diluted with 2.5 mL 1:1 MeCN/H_2O (+0.1% TFA) and purified by RP-HPLC to afford **11** as a white solid (13.5 mg, 27% yield).

HPLC conditions: C8 xBridge column 5 µm OBD, 20-70% MeCN (+0.05% TFA) over 30 min. Flow rate: 16 mL/min. T_R = 13 min.

ESI-MS(+) for peptide **11**. Chemical Formula = $C_{344}H_{539}N_{93}O_{108}S_6$. MW = 7898.00 g/mol. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1975.50, found: 1975.49; $[M+5H^+]^{5+}$ *m/z*: 1580.60, found: 1579.94; $[M+6H^+]^{6+}$ *m/z*: 1317.33, found: 1316.89; $[M+7H^+]^{7+}$ *m/z*: 1129.28, found: 1128.85.; $[M+8H^+]^{8+}$ *m/z*: 988.25, found: 988.00.; $[M+9H^+]^{9+}$ *m/z*: 878.55, found: 878.42; $[M+10H^+]^{10+}$ *m/z*: 790.80, found: 790.68; $[M+11H^+]^{11+}$ *m/z*: 719.00, found: 718.60.

NCL and Thz removal to synthesize KRas[51-166] (12)

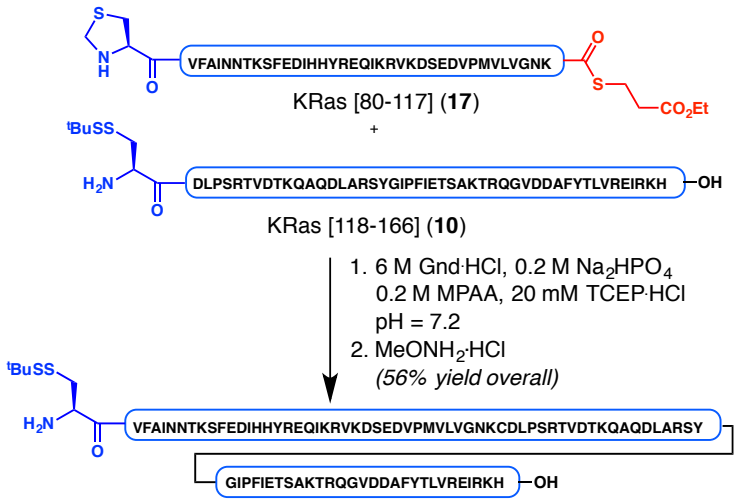


KRas[51-117] (**11**) (10.7 mg, 1.35 µmol) and KRas[118-166] (**10**) (9.3 mg, 1.63 µmol) were weighed into a 1-dram glass vial, and NCL buffer was added (675 µL). The reaction was stirred under argon for 5 hours. At this point a solution of $MeONH_2{\cdot}HCl$ (48 mg, 0.57 mmol) dissolved in NCL buffer (610 µL) was added to the reaction, and the mixture was further stirred for 4 hours, at which point all of the thiazolidine had been removed as indicated by UPLC/MS analysis. The reaction was diluted with 2 mL 1:1 MeCN/H_2O (+0.1% TFA) and purified by RP-HPLC to afford **12** as a white solid (8.4 mg, 46% yield).

HPLC conditions: C4 microsorb column, 30-65% MeCN (+0.05% TFA) over 30 min. *Flow rate:* 20 mL/min. T_R = 14 min.

ESI-MS(+) for peptide **12**. Chemical Formula = $C_{582}H_{921}N_{165}O_{183}S_6$. MW = 13350.10 g/mol. ESI calculated for $[M+7H^+]^{7+}$ *m/z*: 1908.15, found: 1907.98; $[M+8H^+]^{8+}$ *m/z*: 1669.76, found: 1669.35; $[M+9H^+]^{9+}$ *m/z*: 1484.34, found: 1484.23; $[M+10H^+]^{10+}$ *m/z*: 1336.01, found: 1335.79; $[M+11H^+]^{11+}$ *m/z*: 1214.64, found: 1214.36; $[M+12H^+]^{12+}$ *m/z*: 1113.50, found: 1113.18; $[M+13H^+]^{13+}$ *m/z*: 1027.93, found: 1027.67; $[M+14H^+]^{14+}$ *m/z*: 954.58, found: 954.47; $[M+15H^+]^{15+}$ *m/z*: 891.00, found: 890.95; $[M+16H^+]^{16+}$ *m/z*: 835.38, found: 835.38; $[M+17H^+]^{17+}$ *m/z*: 786.30, found: 786.25; $[M+18H^+]^{18+}$ *m/z*: 742.67, found: 742.75; $[M+19H^+]^{19+}$ *m/z*: 703.64, found: 703.53.

NCL between 7 and 10 to afford KRas[80-166] (18):

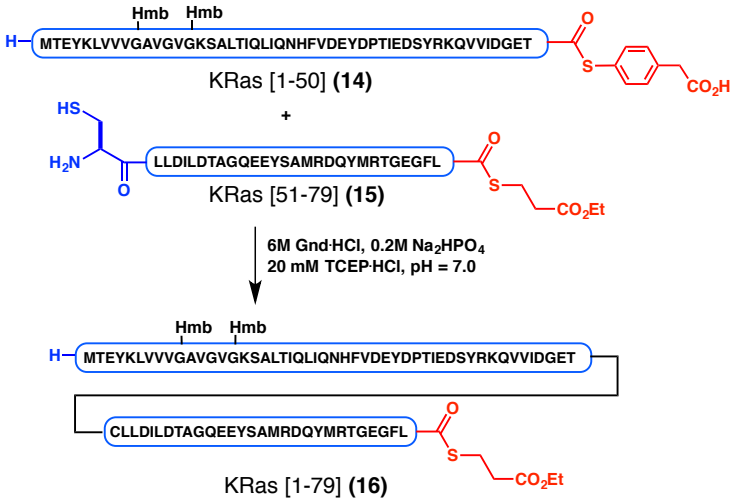


KRas[80-117] (**17**) (11.5 mg, 2.5 µmol) and KRas[118-166] (**10**) (9.3 mg, 1.63 µmol) were weighed into a 1-dram scintillation vial. The peptides were solubilized with NCL Buffer (850 µL, pH ~ 7.2) and stirred under argon for 5 hours. At this point, UPLC/MS analysis indicated full conversion. To this solution was added methoxylamine hydrochloride (66 mg, 0.79 mmol) as a solution in NCL buffer (850 µL). The reaction was stirred a further 7 hours, and was diluted with 2 mL 1:1 MeCN/H_2O (+0.1% TFA). Purification by RP-HPLC afforded KRas[80-166] (**18**) as a white solid (9.2 mg, 56% yield).

HPLC Conditions: C8 xBridge column 5µm OBD, 15-50% MeCN (+0.05% TFA) over 30 min. *Flow rate*: 16 mL/min. T_R = 18 min.

ESI-MS(+) for peptide **18**. Chemical Formula = $C_{440}H_{702}N_{128}O_{135}S_3$. MW = 10041.40 g/mol. ESI calculated for $[M+6H^+]^{6+}$ *m/z*: 1674.56, found: 1674.37; $[M+7H^+]^{7+}$ *m/z*: 1435.49, found: 1435.18; $[M+8H^+]^{8+}$ *m/z*: 1256.18, found: 1255.83; $[M+9H^+]^{9+}$ *m/z*: 1116.71, found: 1116.33; $[M+10H^+]^{10+}$ *m/z*: 1005.14, found: 1004.95; $[M+11H^+]^{11+}$ *m/z*: 913.85, found: 913.75; $[M+12H^+]^{12+}$ *m/z*: 837.78, found: 837.63; $[M+13H^+]^{13+}$ *m/z*: 773.41, found: 773.20; $[M+14H^+]^{14+}$ *m/z*: 718.24, found: 718.08; $[M+15H^+]^{15+}$ *m/z*: 670.43, found: 670.38.

KRas[1-79]-SR (16) synthesis via kinetic chemical ligation between 14 and 15:



MPAA-thioester **14** (all D-residues, 12.8 mg, 2.13 µmol) and alkyl-thioester **15** (all D-residues, 10.8 mg, 3.14 µmol) were weighed into a 1-dram glass scintillation vial equipped with stirbar and screwcap. The fragments were then dissolved in kinetic ligation buffer (1.0 mL, 6M Gnd·HCl, 0.2 M Na_2HPO_4, 20 mM TCEP·HCl, pH = 7.0), and the reaction mixture was stirred under an atmosphere of argon for 7 hours. Ethyl 3-mercaptopropionate (25.0 uL, 0.2 mmol) was added followed by additional TCEP (0.5 M aqueous solution, pH = 7.0, 25 uL, 12.5 µmol). After 30 minutes, the mixture was diluted with 1.0 mL 1:1 MeCN/H_2O (+0.1% TFA) and purified by RP-HPLC. Purest fractions were combined and lyophilized to afford **16** as a white solid (7.1 mg, 36%

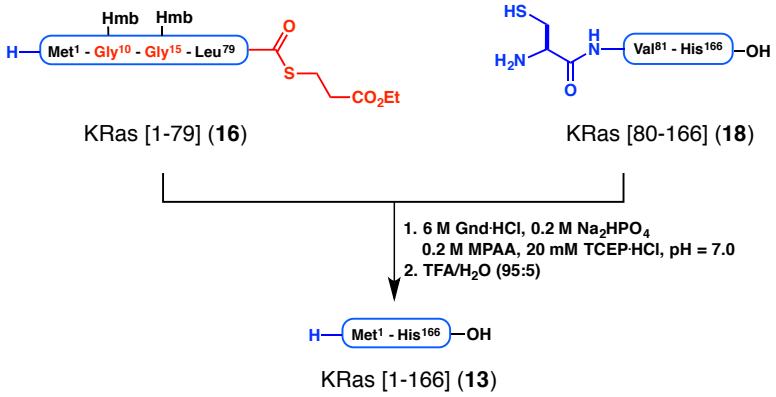
yield). Occasionally, a small amount of the macrocycle derived from KRas[51-79] coeluted with the desired fragment. In this case, the material was used as is in subsequent steps, and this macrocycle had no impact on later steps.

This same reaction was run using fragments derived from all L-residues on a similar scale (2.00 µmol limiting reagent) to obtain the analogous all L peptide (6.0 mg, 32% yield).

HPLC Conditions: xBridge Prep C8 column 5µm OBD (19 x 150 mm), 20-70% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow Rate*: 15 mL/min. T_r = 16.5 minutes.

ESI-MS(+) for peptide **16**: Chemical Formula = $C_{412}H_{639}N_{99}O_{133}S_5$. MW = 9267.50 g/mol. ESI calculated for $[M+5H^+]^{5+}$ *m/z*: 1854.50, found: 1854.14; $[M+6H^+]^{6+}$ *m/z*: 1545.58, found: 1545.29; $[M+7H^+]^{7+}$ *m/z*: 1324.93, found: 1324.54; $[M+8H^+]^{8+}$ *m/z*: 1159.44, found: 1159.15; $[M+9H^+]^{9+}$ *m/z*: 1030.72, found: 1030.52; $[M+10H^+]^{10+}$ *m/z*: 927.75, found: 927.40.

Successful NCL and Hmb removal to produce KRas[1-166] (13)



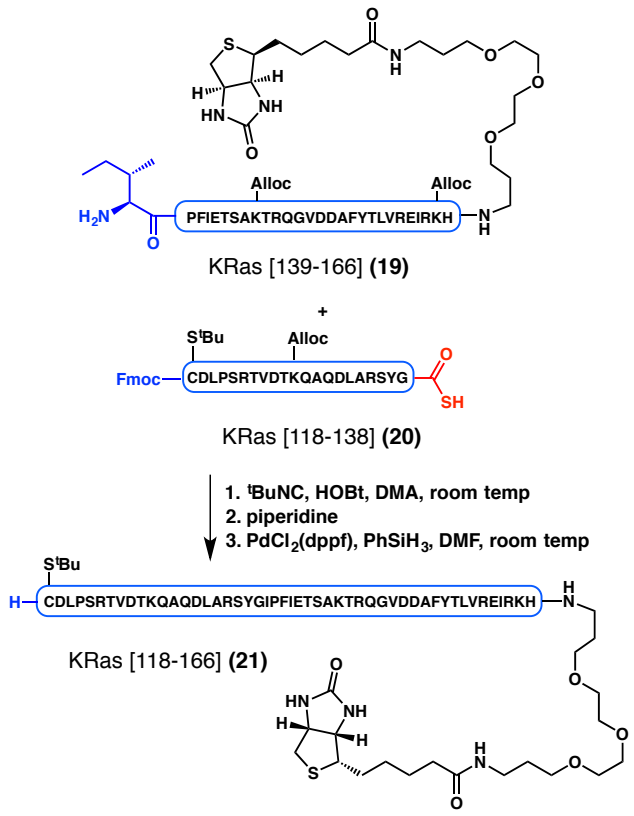
KRas[1-79] **16** (2.80 mg, 0.30 µmol) and KRas[80-166] **18** (3.65 mg, 0.36 µmol) were weighed into a 1-dram scintillation vial and solubilized in NCL buffer (pH ~ 7.2). The reaction was stirred under argon for 11 hours, at which point the mixture was diluted with 0.5 mL 1:1 MeCN/H_2O (+0.1% TFA). The mixture was desalted by passing the solution through a column of sephadex LH-20 with gravity filtration (~5 mL dry adsorbent, swelled in 1:1 MeCN/H_2O containing 0.1% TFA). 3 mL fractions are

collected, and most of the desired compound elutes in fractions 3-5. Some of the ligation product coelutes in later fractions with guanidine salts, which is lyophilized separately, and run through a second column of Sephadex LH-20. All combined fractions containing the desired compound are pooled and lyophilized. The resulting crude peptide is treated with TFA/H_2O/*i*-Pr_3SiH (95:2.5:2.5) for 1 hour 20 min. The reaction mixture is cooled over a bed of dry ice, and the peptide is precipitated upon the addition of cold diethyl ether, and solids are collected via centrifugation. The resulting pellet is dissolved in 4.5 mL 1:1 MeCN/H_2O (+0.1% TFA), shell-frozen, and lyophilized. The solids are dissolved in 300 mL solubilization buffer (6 M Gnd·HCl, 0.2 M Na_2HPO_4, pH = 7.2) and purified by RP-HPLC to afford **13** as a white solid (910 μg, 16% overall yield)

HPLC conditions: C4 Higgins Semi-Preparative column (4.6 x 250 mm), 30-70% MeCN (+0.05% TFA). *Flow rate*: 1.5 mL/min. T_R = 17.5 min.

ESI-MS(+) for peptide **13**. Chemical Formula = $C_{831}H_{1315}N_{227}O_{262}S_7$. MW = 18902.41 g/mol. ESI calculated for $[M+10H^+]^{10+}$ *m/z*: 1891.24, found: 1890.81; $[M+11H^+]^{11+}$ *m/z*: 1719.40, found: 1719.38; $[M+12H^+]^{12+}$ *m/z*:1576.20, found:1576.04; $[M+13H^+]^{13+}$ *m/z*: 1455.03, found: 1454.83; $[M+14H^+]^{14+}$ *m/z*: 1351.17, found: 1351.09; $[M+15H^+]^{15+}$ *m/z*: 1261.16, found: 1260.86; $[M+16H^+]^{16+}$ *m/z*: 1182.40, found: 1182.33; $[M+17H^+]^{17+}$ *m/z*: 1112.90, found: 1112.88; $[M+18H^+]^{18+}$ *m/z*: 1051.13, found: 1050.92; $[M+19H^+]^{19+}$ *m/z*: 995.86, found: 995.57; $[M+20H^+]^{20+}$ *m/z*: 946.12, found: 946.07; $[M+21H^+]^{21+}$ *m/z*: 901.11, found: 901.15; $[M+22H^+]^{22+}$ *m/z*: 860.20, found: 860.12; $[M+23H^+]^{23+}$ *m/z*: 822.84, found: 822.63; $[M+24H^+]^{24+}$ *m/z*: 788.60, found: 788.80.

Isonitrile-Mediated Ligation between 19 and 20 to produce KRas[118-166] (21):



Thioacid 20 (all L-amino acids, 13.3 mg, 4.86 μmol) and biotin-labelled **19** (15.7 mg, 4.03 μmol) were directly weighed into a 1-dram glass scintillation vial, and HOBt (6.7 mg, 50.0 μmol) was added followed by ^{t}BuNC (1.3 mg, 15.5 μmol, as a freshly prepared stock solution in 930 μL anhydrous *N,N*-dimethylacetamide). The mixture was stirred for 48 hours, whereupon piperidine (0.200 mL) was added to remove the N-terminal Fmoc group. After stirring for 30 minutes at room temperature, the reaction was transferred by micropipette into a 15 mL conical centrifuge vial. Cold diethyl ether (12 mL) was added to precipitate the crude peptide, the vial was centrifuged (5 minutes at 5000 rpm, 10^{o}C), and ether was decanted. The residue was redissolved in 1:1 MeCN/H_2O (+0.1% TFA), shell frozen, and lyophilized.

The resulting crude lyophilized peptide was dissolved in DMF (1.22 mL, degassed for 30 minutes with Argon sparge), and $PhSiH_3$ (25.0 μL, 203 μmol) was added followed

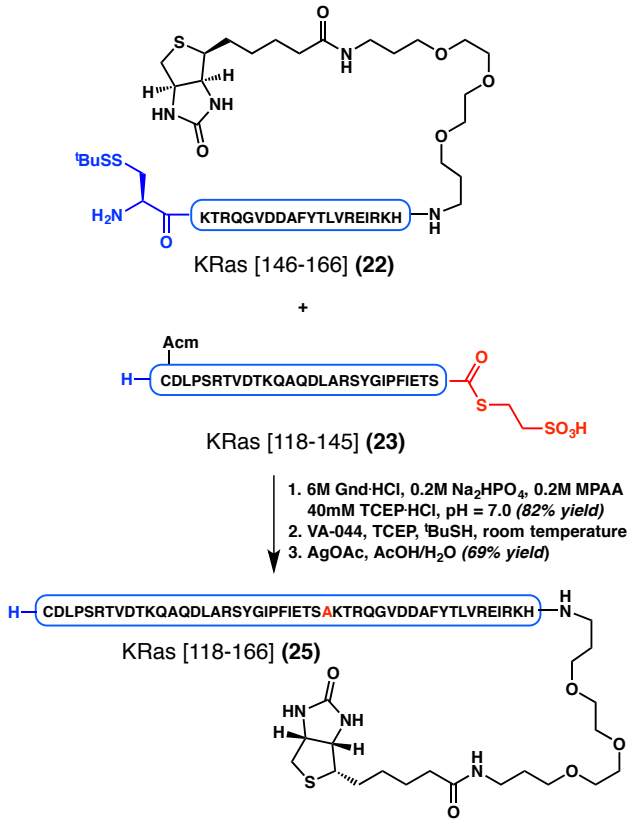
by $PdCl_2$(dppf) (1.50 mg, 2.04 μmol, as a freshly prepared stock solution in 270 μL DMF). The reaction was stirred for 20 minutes under argon at which point UPLC/MS showed full deprotection. After stirring a further 10 minutes, the vial was cooled over a bed of dry ice for 1 minute, and cold ether (10.5 mL) was added to precipitate the peptide. The precipitated peptide was centrifuged, the supernatant was decanted, and the residue was further washed with cold ether (10 mL), and again centrifuged and decanted as before. The resulting crude peptide was dissolved in 4 mL 1:1 MeCN/H_2O (+0.1% TFA), and immediately purified by RP-HPLC to afford **21** as a white solid (13.8 mg, 56% yield).

The analogous reaction run with peptide fragments composed of all-D amino acid residues gave a similar yield on a 5.14 μmol scale (limiting reagent), affording 17.8 mg KRas[118-166] **21** (57% yield).

HPLC Conditions: xBridge Prep C8 column 5μm OBD (19 x 150 mm), 20-60% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow Rate*: 15 mL/min. T_r = 15.4 minutes.

ESI-MS for peptide **21**: Chemical Formula = $C_{268}H_{436}N_{76}O_{81}S_3$. MW = 6115.07 g/mol. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1529.76, Found: 1529.61; $[M+5H^+]^{5+}$ *m/z*: 1224.01, found: 1223.95; $[M+6H^+]^{6+}$ *m/z*: 1020.17, found: 1019.96; $[M+7H^+]^{7+}$ *m/z*: 874.58, found: 874.42; $[M+8H^+]^{8+}$ *m/z*: 765.38, found: 765.26; $[M+9H^+]^{9+}$ *m/z*: 680.45, found: 680.42

NCL between 22 and 23; Desulfurization of 24 and Acm Removal to produce 25:



Thioester **23** (all L-amino acids, 4.32 mg, 1.30 µmol) and **22** (all L-amino acids, 4.40 mg, 1.44 µmol) were each weighed directly into a 2.0 mL eppendorf plastic tube. The solids were solubilized in degassed NCL buffer (650 µL, 6M Gnd·HCl, 0.2 M Na_2HPO_4, 0.2M MPAA, 40 mM TCEP·HCl, pH = 7.0). The vial was vortexed and centrifuged briefly, then stirred under argon at room temperature. After 6 hours, the reaction was quenched by the addition of 1:1 MeCN/H_2O (0.1% TFA). The solution was purified by RP-HPLC. Purest fractions were combined, shell-frozen, and lyophilized to afford **24** as a white solid (6.60 mg, 82% yield).

HPLC Conditions: xBridge Prep C8 column 5µm OBD (19 x 150 mm), 15-65% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow Rate*: 16 mL/min. T_r = 14.5 minutes.

ESI-MS(+) for peptide **24**. Chemical Formula = $C_{267}H_{433}N_{77}O_{82}S_3$. MW = 6130.04 g/mol. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1533.51, found: 1533.43; $[M+5H^+]^{5+}$ *m/z*: 1227.00, found: 1226.93; $[M+6H^+]^{6+}$ *m/z*: 1022.67, found: 1022.58; $[M+7H^+]^{7+}$ *m/z*: 876.72, found: 876.70; $[M+8H^+]^{8+}$ *m/z*: 767.25, found: 767.21.

Desulfurization of 24: The ligation product **24** (5.58 mg, 0.89 μmol) was weighed into a 2.0 mL eppendorf plastic tube, and was solubilized in 0.45 mL desulfurization buffer (6M Gnd·HCl, 0.2 M Na_2HPO_4, pH not adjusted, degassed with argon bubbling and sonication for 30 minutes). To this solution was added TCEP (0.5 M solution, pH = 7.0, 0.45 mL, 225 μmol) and ^{t}BuSH (74 μL, 0.656 mmol), followed by VA-044 (freshly prepared 0.1M aqueous solution, 225 μL, 22.5 μmol). The reaction was then stirred under argon at room temperature for 4.5 hours. At this point, the desulfurization was determined to be complete by UPLC/MS analysis. The mixture was desalted by passing the solution through two stacked solid phase extraction cartidges (Seppak C18 plus, 360 mg adsorbent). The product was eluted, collecting 2 mL fractions, with a stepwise gradient of 0-80% MeCN in H_2O (+0.05% TFA), increasing the percentage of MeCN by 10% every two fractions. Guanidine salts elute in fractions 1-3, and the product elutes in fractions 8-13, coeluting with ^{t}BuSH in some fractions. All fractions containing the desulfurization product were combined and diluted with H_2O, shell-frozen, and lyophilized.

ESI-MS(+) for **24 (Desulfurized)**: Chemical Formula = $C_{267}H_{433}N_{77}O_{82}S_2$. MW = 6097.98 g/mol. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1525.49, found: 1525.62; $[M+5H^+]^{5+}$ *m/z*: 1220.60, found: 1220.63; $[M+6H^+]^{6+}$ *m/z*: 1017.33, found: 1017.33; $[M+7H^+]^{7+}$ *m/z*: 872.14, found: 872.12; $[M+8H^+]^{8+}$ *m/z*: 763.25, found: 763.31; $[M+9H^+]^{9+}$ *m/z*: 678.55, found: 678.89

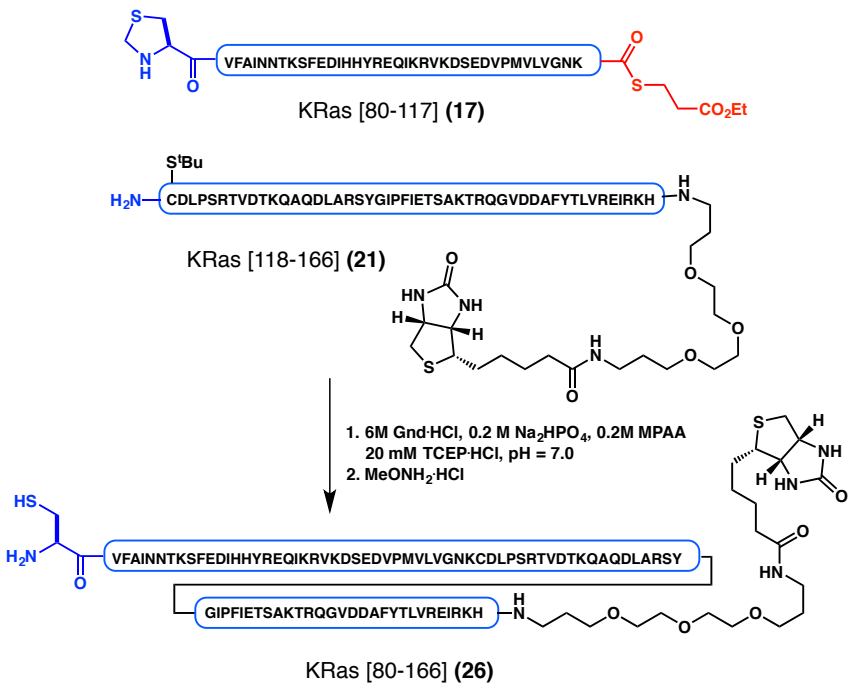
To this lyophilized material (assumed 0.89 μmol) in a 15 mL plastic centrifuge tube was added AgOAc (14.4 mg, 86 μmol) followed by 70% aqueous AcOH (1.32 mL, degassed with argon sparge and sonication for 15 minutes). The resulting suspension was agitated at room temperature under argon. After 3 hours of stirring, UPLC/MS analysis indicated

quantitative deprotection. To this reaction was added 1,2-dithiothreitol (28.0 mg, 0.182 mmol, as a solution in 0.4 mL 50% aqueous AcOH). The suspension was then agitated for 2 hours at room temperature, at which point the vial was centrifuged and the supernatant was collected. The solid pellet was then washed twice with 0.8 mL 50% aqueous AcOH, with centrifugation and collection of the supernatant. The combined supernatant solutions were concentrated partway with N_2 stream, diluted with H_2O, shell-frozen and lyophilized. The solid residue was dissolved in 1.5 mL 1:1 MeCN/H_2O (0.1% TFA) and purified by RP-HPLC. All fractions containing the desired peptide were combined, shell-frozen, and lyophilized to afford **25** as a white solid (3.74 mg, 69% yield over 2 steps, 57% yield overall)

HPLC Conditions: xBridge Prep C8 column 5μm OBD (19 x 150 mm), 15-65% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow Rate*: 16 mL/min. T_r = 14.5 minutes.

ESI-MS(+) for peptide **25**: Chemical Formula = $C_{264}H_{428}N_{76}O_{81}S_2$. MW = 6026.90 g/mol. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1507.72, found: 1507.61; $[M+5H^+]^{5+}$ *m/z*: 1206.38, found: 1206.37; $[M+6H^+]^{6+}$ *m/z*: 1005.48, found: 1005.40; $[M+7H^+]^{7+}$ *m/z*: 861.98, found: 861.92; $[M+8H^+]^{8+}$ *m/z*: 754.36, found: 754.31; $[M+9H^+]^{9+}$ *m/z*: 670.65, found: 670.71.

NCL between 17 and 21 and Thz removal to produce KRas[80-166] (26):



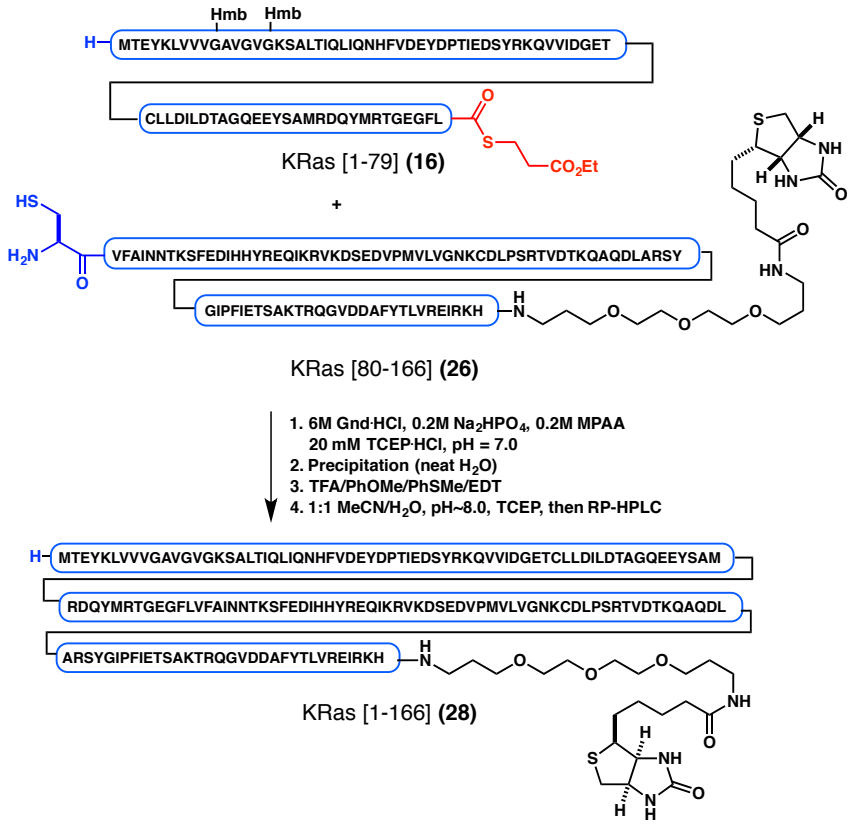
Thioester **17** (all D-residues, 10.32 mg, 2.20 μmol) and **21** (all D-residues, 10.83 mg, 1.77 μmol) were directly weighed into a 2 mL plastic eppendorf tube with a stirbar. The fragments were dissolved in NCL buffer (0.87 mL, pH = 7.0) and stirred under argon for 5 hours, at which point the reaction was complete by UPLC/MS analysis. Methoxylamine hydrochloride (33 mg, 0.40 mmol) was added directly to the reaction, the reaction mixture carefully adjusted to pH~4.5 with 6M HCl, and the reaction stirred a further 4.5 hours at room temperature, at which point all of the thiazolidine had been removed as judged by UPLC/MS analysis. The reaction was diluted with 0.6 mL 1:1 MeCN/H_2O (+0.1% TFA) and purified by RP-HPLC. Combination of the purest fractions and lyophilization afforded **26** as a white solid (11.73 mg, 63% yield).

The analogous reaction was run using peptides consisting of all L amino acids on a 2.25 μmol scale (limiting reagent) to yield the all-L residue sequence (15.6 mg, 66% yield).

HPLC Conditions: xBridge Prep C8 column 5µm OBD (19 x 150 mm), 15-50% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow Rate*: 16 mL/min. T_r = 16 minutes.

ESI-MS(+) for peptide **26**: Chemical Formula = $C_{460}H_{738}N_{132}O_{139}S_4$. MW = 10469.99 g/mol. ESI calculated for $[M+6H^+]^{6+}$ *m/z*: 1745.99, found: 1745.73; $[M+7H^+]^{7+}$ *m/z*: 1496.71, found: 1496.13; $[M+8H^+]^{8+}$ *m/z*: 1309.75, found: 1309.56; $[M+9H^+]^{9+}$ *m/z*: 1164.99, found: 1164.04; $[M+10H^+]^{10+}$ *m/z*: 1047.99, found: 1047.50; $[M+11H^+]^{11+}$ *m/z*: 952.81, found: 952.57; $[M+12H^+]^{12+}$ *m/z*: 873.50, found: 873.47; $[M+13H^+]^{13+}$ *m/z*: 806.38, found: 806.46; $[M+14H^+]^{14+}$ *m/z*: 748.86, found: 748.68; $[M+15H^+]^{15+}$ *m/z*: 699.00, found: 699.30.

Synthesis of biotinylated KRas[1-166] (28)



Hmb-derivative **16** (all D-amino acids, 1.66 mg, 0.185 µmol) and **26** (all D-amino acids, 3.64 mg, 0.347 µmol) were each directly weighed into a 1.5 mL eppendorf tube containing a stirbar. The fragments were dissolved in NCL buffer (206 uL, pH = 7.0) and additional TCEP was added (0.5 M aqueous solution, pH = 7.0, 9 uL, 4.5 µmol).

The reaction was stirred under argon for 6.5 hours, at which point the reaction mixture was transferred via micropipette into a 14 mL conical centrifuge vial. The solution transfer was quantitated with an additional 100 μL fresh NCL buffer. The crude peptide was then directly precipitated by the addition of cold H_2O (3 mL). The mixture was then frozen in liquid nitrogen, and allowed to thaw with stirring. The vial was then centrifuged (5000 rpm, 5 min, 10°C), and the supernatant was carefully collected via syringe. The resulting solid residue was further washed with H_2O (2 mL), refrozen and thawed as before, and then centrifuged and supernatant removed as before. The resulting white solid pellet was dissolved in 5 mL 1:1 MeCN/H_2O (+0.1% TFA), shell frozen and lyophilized. Analysis of the combined aqueous supernatants and the redissolved pellet showed that most of **26** remained in the mother liquor with only trace **27** remaining dissolved, while the redissolved pellet consisted almost exclusively of **27**. The crude lyophilized peptide was then subjected to cocktail R deprotection (90/5/3/2 TFA/thioanisole/EDT/anisole, 450 uL) for 2 hours under argon. The crude peptide was directly precipitated upon the addition of cold ether (5.5 mL). After cooling on a bed of dry ice for 5 minutes, the vial was centrifuged, ether decanted, and the resulting solid was dissolved in 5 mL 1:1 MeCN/H_2O (+0.1% TFA) and lyophilized. The residue was solubilized in 1.0 mL 1:1 MeCN/H_2O (+0.1% TFA) with sonication, pH adjusted to ~8.0 using 3N NaOH, and TCEP was added (0.5 M aqueous solution, pH = 7.0, 30 uL, 15 μmol). This solution was then purified by RP-HPLC. Purest fractions were combined and lyophilized to yield deprotected biotinylated all-D residue KRas [1-166] **28** as a white solid (0.92 mg, 25% yield overall).

The analogous reaction was run using peptides consisting of all-L residues (0.17 μmol limiting reagent). The overall yield was 0.73 mg (21% yield).

HPLC Conditions: xBridge Prep C4 column 5μm OBD (19 x 150 mm), 20-70% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow Rate*: 16 mL/min. T_r = 14.5 minutes.

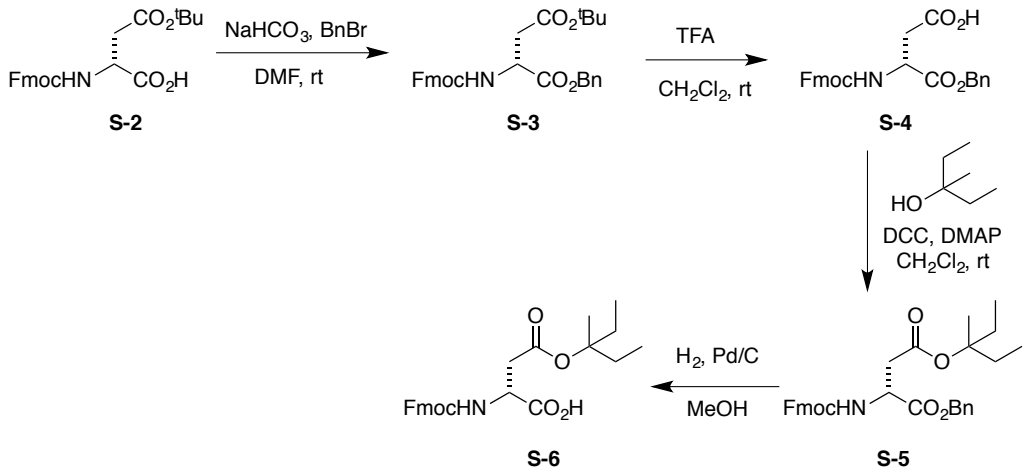
ESI-MS(+) for Hmb-protected KRas[1-166] **27**: Chemical Formula = $C_{867}H_{1367}N_{231}O_{270}S_8$. MW = 19603.30 g/mol. ESI calculated for $[M+11H^+]^{11+}$ *m/z:* 1783.11*,* found: 1782.89; $[M+12H^+]^{12+}$ *m/z*: 1634.60, found: 1634.42; $[M+13H^+]^{13+}$ *m/z*:

1508.94, found: 1508.98; $[M+14H^+]^{14+}$ *m/z*: 1401.24, found: 1401.16; $[M+15H^+]^{15+}$ *m/z*: 1307.89, found: 1307.90; $[M+16H^+]^{16+}$ *m/z*: 1226.21, found: 1226.35; $[M+17H^+]^{17+}$ *m/z*: 1154.13, found: 1153.95; $[M+18H^+]^{18+}$ *m/z*: 1090.07, found: 1090.03; $[M+19H^+]^{19+}$ *m/z*: 1032.75, found: 1032.64; $[M+20H^+]^{20+}$ *m/z*: 981.16, found: 981.10; $[M+21H^+]^{21+}$ *m/z*: 934.49, found: 934.36; $[M+22H^+]^{22+}$ *m/z*: 892.06, found: 891.97; $[M+23H^+]^{23+}$ *m/z*: 853.32, found: 853.34; $[M+24H^+]^{24+}$ *m/z*: 817.80, found: 817.78.

ESI-MS(+) for purified KRas[1-166] **28**: Chemical Formula = $C_{851}H_{1351}N_{231}O_{266}S_8$. MW = 19331.00 g/mol. ESI calculated for $[M+11H^+]^{11+}$ *m/z:* 1758.35*,* found: 1758.26; $[M+12H^+]^{12+}$ *m/z*: 1611.91, found: 1611.85; $[M+13H^+]^{13+}$ *m/z*: 1488.00, found: 1487.72; $[M+14H^+]^{14+}$ *m/z*: 1381.79, found: 1381.68; $[M+15H^+]^{15+}$ *m/z*: 1289.73, found: 1289.60; $[M+16H^+]^{16+}$ *m/z*: 1209.19, found: 1209.22; $[M+17H^+]^{17+}$ *m/z*: 1138.11, found: 1138.00; $[M+18H^+]^{18+}$ *m/z*: 1074.94, found: 1074.89; $[M+19H^+]^{19+}$ *m/z*: 1018.42, found: 1018.31; $[M+20H^+]^{20+}$ *m/z*: 967.55, found: 967.50; $[M+21H^+]^{21+}$ *m/z*: 921.52, found: 921.50; $[M+22H^+]^{22+}$ *m/z*: 879.68, found: 879.63; $[M+23H^+]^{23+}$ *m/z*: 841.48, found: 841.43; $[M+24H^+]^{24+}$ *m/z*: 806.46, found: 806.46.

Part V. Preparation of specialized D-amino acid residues, α-thioester residues, and nucleotides

Synthesis of Fmoc-D-Asp(Mpe)-OH:



S-3: To a solution of Fmoc-D-Asp(^{t}Bu)-OH (**S-2**) (10.0 g, 24.3 mmol) in DMF (125 mL) was added $NaHCO_3$ (5.13 g, 61 mmol) followed by BnBr (8.70 mL, 73.2 mmol). The reaction was stirred at room temperature for 48 hours, at which point it was diluted with EtOAc (150 mL), and the mixture washed with brine (200 mL). The aqueous layer was reextracted once with EtOAc (150 mL). Combined organics were washed again with brine (4 x 100 mL), dried ($MgSO_4$), and concentrated. Purification by silica gel chromatography (100% Hexane → 9:1 Hexane:EtOAc → 2:1 Hexane:EtOAc) afforded Fmoc-D-Asp(^{t}Bu)-OBn **S-3** as a white solid (11.10 g, 91% yield). Spectroscopic data matched that previously reported.[46]

$[\alpha]_D^{22.1} = -15.8^o$ (c = 0.44, $CHCl_3$)

Fmoc-D-Asp(^{t}Bu)-OBn **S-3**: ^{1}H NMR (600 MHz, $CDCl_3$) 7.76 (d, J = 7.5 Hz, 2H), 7.59 (d, J = 7.1 Hz, 2 H), 7.40 (t, J = 7.5 Hz, 2 H), 7.28-7.34 (m, 7H), 5.83 (d, J = 8.6 Hz, 1H), 5.23 (d, J = 12.2 Hz, 1H), 5.17 (d, J = 12.2 Hz, 1 H), 4.64 (m, 1H), 4.41 (dd, J = 10.4, 7.4 Hz, 1H), 4.33 (dd, J = 10.4, 7.6 Hz, 1H), 4.23 (app t, J = 7.3 Hz, 1H), 2.97 (dd, J = 17.0, 4.5 Hz, 1H), 2.78 (dd, J = 17.0, 4.4 Hz, 1H), 1.41 (s, 9H). ^{13}C NMR (150 MHz, $CDCl_3$) 171.0, 170.2, 156.2, 144.1, 143.9, 141.5, 141.4, 135.5, 128.8, 128.6,

128.5, 127.9, 127.3, 125.4, 125.3, 120.2, 82.1, 67.7, 67.5, 50.9, 47.3, 38.0, 28.2. LRMS ESI(+)-MS for $C_{30}H_{31}NO_6$ *m/z:* 524.25 $[M+Na^+]^+$.

To a solution of **S-3** (5.52 g, 11.01 mmol) in CH_2Cl_2 (110 mL) was added TFA (8.42 mL, 110 mmol). The solution was left to stir overnight, at which point the solvent was removed on a rotatory evaporator. The crude material was dissolved in a minimal amount of Hexane/EtOAc with gentle heating, and left to precipitate in a freezer overnight. Solvent was decanted, and solids were washed with cold hexanes and dried under high vacuum to afford **S-4** as a white solid (3.64 g, 74% yield). Spectroscopic data matched previously reported.[46]

$[\alpha]_D^{21.1}$ = -10.2° (*c* = 0.73, $CHCl_3$), lit. $[\alpha]_D^{21}$ = +12° for enantiomer (c = 1.84, $CHCl_3$)

Fmoc-D-Asp-OBn **S-4**: ^{1}H NMR (600 MHz, $CDCl_3$) 7.75 (d, *J* = 7.5 Hz, 2H), 7.58 (m, 2H), 7.39 (t, *J* = 7.5 Hz, 1H), 7.27-7.40 (m, 7 H), 5.82 (d, *J* = 8.3 Hz, 1H), 5.21 (m, 2 H), 4.71 (m, 1H), 4.43 (dd, *J* = 10.4, 7.5 Hz, 1H), 4.36 (dd, *J* = 10.4, 7.3 Hz, 1H), 4.21 (app t, *J* = 7.1 Hz, 1H), 3.11 (dd, *J* = 17.5, 4.1 Hz, 1H), 2.94 (dd, *J* = 17.5, 4.0 Hz, 1H). ^{13}C NMR (150 MHz, $CDCl_3$) 175.6, 170.6, 156.2, 144.0, 143.8, 141.5 (2 C), 135.2, 128.8, 128.7, 128.5, 128.0, 127.3, 125.3, 120.2, 68.0, 67.6, 50.5, 47.3, 36.5. LRMS: ESI(+)-MS for $C_{26}H_{23}NO_6$ *m/z:* 468.13 $[M+Na^+]^+$.

To a 0°C solution of *N,N'*-dicyclohexylcarbodiimide (1.297g, 6.3 mmol) and 3-methyl-3-pentanol (0.912 mL, 7.4 mmol) in anhydrous CH_2Cl_2 (15 mL) was added a solution of Fmoc-D-Asp-OBn (**S-4**) (2.50 g, 5.68 mmol) and 4-dimethylaminopyridine (170 mg, 1.40 mmol) in CH_2Cl_2 (10 mL) over ca. 3-4 minutes. The reaction was stirred at 0 °C for 10 minutes, then warmed to room temperature for 8 hours. The mixture becomes very thick with white solids, which is difficult to stir, and this solid breaks up as the reaction proceeds until it becomes a finer brownish suspension. Solids were filtered through celite and washed with CH_2Cl_2 (70 mL). The filtrate was washed with sat. aq. NH_4Cl (30 mL) followed by brine (30 mL). The organic layer was dried ($MgSO_4$) and concentrated. Purification by column chromatography (9:1 Hexane/EtOAc → 4:1 Hexane/EtOAc) yielded Fmoc-Asp(Mpe)-OBn **S-5** as a white solid (1.457g, 48% yield). ^{1}H and ^{13}C NMR data matched that previously reported.[47]

$[\alpha]_D^{20.5}$ = -21.1° (*c* = 0.36, $CHCl_3$) Lit. $[\alpha]_D$ = +17° (*c* = 1.0, $CHCl_3$) for Fmoc-L-Asp(Mpe)-OBn

Fmoc-Asp(Mpe)-OBn **S-5**: ^{1}H NMR (600 MHz, $CDCl_3$) 7.76 (d, *J* = 7.5 Hz, 2 H), 7.59 (m, 2 H), 7.39 (app t, *J* = 7.5 Hz, 2 H), 7.28-7.34 (m, 7 H), 5.85 (d, *J* = 8.6 Hz, 1 H), 5.21 (m, 2 H), 4.65 (m, 1 H), 4.41 (dd, *J* = 10.3, 7.5 Hz, 1 H), 4.32 (dd, *J* = 10.3, 7.5 Hz, 1 H), 4.22 (app t, *J* = 7.2 Hz, 1 H), 3.00 (dd, *J* = 17.1, 4.3 Hz, 1 H), 2.81 (dd, *J* = 17.1, 4.2 Hz, 1 H), 1.83 (m, 2 H), 1.72 (m, 2 H), 1.34 (s, 3 H), 0.82 (t, *J* = 7.5 Hz, 3 H). ^{13}C NMR (125 MHz, $CDCl_3$) 171.0, 170.1, 156.2, 144.1, 143.9, 141.5, 135.4, 128.8, 128.6, 128.5, 127.9, 127.2, 125.4, 125.3, 120.2, 87.5, 67.7, 67.5, 50.8, 47.3, 37.7, 30.5 (2 C), 23.0, 8.2. LRMS: ESI(+)-MS for $C_{32}H_{35}NO_6$ *m/z:* 552.32 $[M+Na^+]^+$.

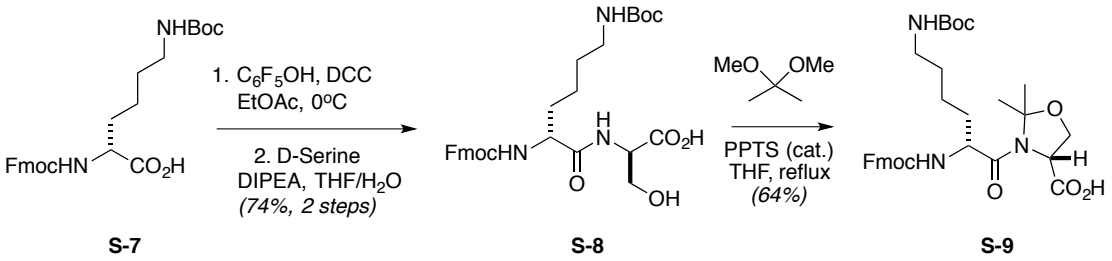
Fmoc-D-Asp(Mpe)-OBn **S-5** (1.415g, 2.67 mmol) was dissolved in absolute EtOH (110 mL), and degassed with argon sparge for 10 minutes. Palladium on carbon (10%, 212 mg, 0.20 mmol) was added, and hydrogen gas was bubbled from a balloon into the reaction mixture for 20 minutes, at which point TLC analysis indicated consumption of starting material. The reaction was filtered through celite, which was further washed with EtOAc (75 mL), and the filtrate was concentrated. Purification by silica gel chromatography (9:1 Hexane:EtOAc +1% AcOH → 4:1 Hexane:EtOAc +1% AcOH), followed by lyophilization twice from benzene afforded Fmoc-D-Asp(Mpe)-OH **S-6** as a white solid (955 mg, 81% yield). ^{1}H and ^{13}C NMR data matched that previously reported.[47]

$[\alpha]_D^{22.5}$ = -30.8° (*c* = 0.45, $CHCl_3$) Lit. $[\alpha]_\mathbf{D}$ = +31.0° (*c* = 1.0, $CHCl_3$) for Fmoc-L-Asp(Mpe)-OH

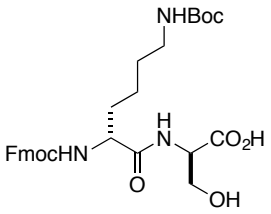
Fmoc-D-Asp(Mpe)-OH **S-6:** ^{1}H NMR (600 MHz, $CDCl_3$) 7.76 (d, *J* = 7.6 Hz, 2H), 7.60 (t, *J* = 6.8 Hz, 2H), 7.40 (t, *J* = 7.6 Hz, 2H), 7.31 (t, *J* = 7.6 Hz, 2H), 5.84 (d, *J* = 8.5 Hz, 1H), 4.65 (m, 1H), 4.43 (dd, *J* = 10.6, 7.2 Hz, 1H), 4.36 (dd, *J* = 10.6, 7.2 Hz, 1H), 4.24 (t, *J* = 7.2 Hz, 1H), 3.04 (dd, *J* = 17.3, 4.1 Hz, 1H), 2.83 (dd, *J* = 17.3, 4.9 Hz, 1H), 1.88 (m, 2H), 1.65 (m, 2H), 1.38 (s, 3H), 0.85 (t, *J* = 7.5 Hz, 6H). ^{13}C NMR (150 MHz, $CDCl_3$) 175.3, 170.5, 156.3, 144.0, 143.9, 141.5, 128.0, 127.3, 125.4, 125.3, 120.2,

88.1, 67.6, 50.5, 47.3, 37.6, 30.6, 30.5, 23.0, 8.2. LRMS: ESI(+)-MS for $C_{25}H_{29}NO_6$ *m/z:* 462.84 $[M+Na^+]^+$.

General Procedure for Synthesis of all-D Pseudoproline Dipeptides:[48,49]



S-7 **S-8** **S-9**



S-8

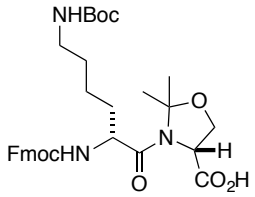
Fmoc-D-Lys(Boc)-OH (**S-7**) (2.81 g, 6.00 mmol) and pentafluorophenol (1.14 g, 6.20 mmol) were suspended in EtOAc (66 mL), and the mixture was cooled in a 0°C ice bath. To the rapidly stirred mixture was added *N,N'*-dicylohexylcarbodiimide (1.36 g, 6.60 mmol), and the reaction was stirred at this temperature for 3 hours. The reaction mixture was then warmed to room temperature and filtered through a pad of celite. The solids were washed with additional EtOAc (20 mL), and the combined filtrate was washed with sat. aq. $NaHCO_3$ (30 mL) followed by brine (30 mL). The organic layer was dried ($MgSO_4$) and concentrated to afford a white solid residue, which was used in the next step without purification.

The resulting crude pentafluorophenol ester (assumed 6.0 mmol) was dissolved in THF (30 mL), and D-Serine was added (1.26 g, 12.0 mmol, dissolved in 12 mL H_2O). *N,N*-Diisopropylethylamine was then added (2.09 mL, 12.0 mmol), and the reaction was stirred for 25 minutes, at which point UPLC/MS analysis indicated full conversion. After 10 minutes longer, the solvent was removed *in vacuo*, and the residue was taken

up in EtOAc (50 mL). The organic phase was washed with 1M HCl (30 mL), and the aqueous layer was further extracted with EtOAc (2 x 50 mL). The combined organic layers were washed with brine (30 mL), dried ($MgSO_4$), and concentrated. Purification by silica gel chromatography (1→10% MeOH in CH_2Cl_2) afforded Fmoc-D-Lys(Boc)-D-Ser-OH **S-8** as a white solid (2.465 g, 74% yield over 2 steps).

$[\alpha]_D^{21.7} = -13.2$ (c = 0.3, $CHCl_3$)

S-8: ^{1}H NMR (600 MHz, MeOH-d_4) 7.78 (d, J = 7.4 Hz, 2H), 7.65 (m, 2 H), 7.38 (t, J = 7.4 Hz, 2H), 7.30 (J = 7.4 Hz, 2H), 4.47 (m, 1H), 4.37 (d, J = 6.9 Hz), 4.22 (m, 1H), 4.15 (m, 1H), 3.90 (dd, J = 11.1, 4.4 Hz, 1H), 3.81 (dd, J = 11.0, 3.4 Hz, 1H), 3.02 (m, 1H), 1.81 (m, 1H), 1.65 (m, 1H), 1.41-1.50 (m, 13 H). ^{13}C NMR (150 MHz, MeOH-d_4) 174.9, 158.6, 145.4, 145.2, 142.6, 128.8, 128.2, 126.3, 126.2, 120.9, 79.9, 68.0, 62.9, 56.4, 56.1, 41.1, 33.0, 30.6, 28.8, 24.1. LRMS: ESI-(+) for $C_{29}H_{37}N_3O_8$ *m/z:* 556.48 $[M+H^+]^+$.

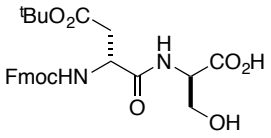


S-9

Dipeptide **S-8** (1.896 g, 3.41 mmol) was dissolved in anhydrous THF (69.0 mL), and pyridinium *p*-toluenesulfonate (171.0 mg, 0.68 mmol) was added, followed by 2,2-dimethoxypropane (2.77 mL, 22.6 mmol). The resulting solution was heated to reflux for 9 hours, passing the condensate over 4Å molecular sieves. At this point the reaction was quenched by the addition of triethylamine (0.15 mL, 1.08 mmol), and the solvent was removed under reduced pressure. Purification by silica gel chromatography (1→10% MeOH in CH_2Cl_2) afforded the pseudoproline dipeptide **S-9** as a white solid (1.290 g, 64% yield). Spectroscopic data closely matched that of the commercially available all L-residue form of this pseudoproline.

$[\alpha]_D^{21.5} = +33.3$ (c = 0.40, $CHCl_3$)

Fmoc-D-Lys(Boc)-D-Ser($\psi^{Me,Me}$Pro)-OH **S-9**: ^{1}H NMR (600 MHz, DMSO-d_6) Rotameric mixture 7.89 (d, J = 7.5 Hz, 2 H), 7.74 (m, 2 H), 7.41 (t, J = 7.4 Hz, 2 H), 7.32 (m, 2 H), 6.76 (m, 1 H), 4.62 (m, 1 H), 4.01-4.29 (m, 6 H), 2.89 (m, 2 H), 1.20-1.58 (m, 21 H). ^{13}C NMR (150 MHz, DMSO-d_6) 171.7, 171.4, 169.6, 169.0, 156.0, 155.6, 155.4, 155.3, 144.0, 143.7, 140.7, 140.6, 127.6, 127.0, 125.4, 120.0, 95.3, 77.3, 66.9, 65.7, 58.5, 53.2, 46.6, 31.9, 29.2, 28.2, 27.7, 25.8, 25.3, 23.0, 22.5. LRMS: ESI-(+) for $C_{32}H_{41}N_3O_8$ *m/z:* 596.46 $[M+H^+]^+$.

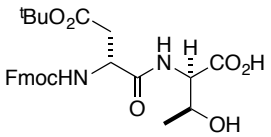


S-10

S-10: Prepared according to an analogous procedure for the synthesis of **S-8**, starting with Fmoc-D-Asp(^{t}Bu)-OH (2.46 g, 6.00 mmol). D-Serine was used for dipeptide formation. Obtained Fmoc-D-Asp(tBu)-D-Ser-OH **S-10** as a white solid (2.63 g, 88% over 2 steps)

$[\alpha]_D^{22.0}$ = -33.5 (c = 0.705, $CHCl_3$)

Fmoc-D-Asp(tBu)-D-Ser-OH **S-10**: ^{1}H NMR (600 MHz, $CDCl_3$) Rotameric Mixture 7.73 (d, J = 7.0 Hz, 2 H), 7.45-7.60 (m, 3 H), 7.36 (m, 2 H), 7.27 (t, J = 7.2 Hz, 2 H), 6.07 (d, J = 6.4 Hz, 1 H), 4.62 (m, 2 H), 4.36 (m, 2 H), 4.18 (t, J = 6.9 Hz, 1 H), 4.04 (m, 1 H), 3.92 (d, J = 11.2 Hz), 2.80 (dd, J = 15.4, 4.6 Hz, 1 H), 2.71 (dd, J = 15.4, 4.6 Hz, 1 H), 1.41 (s, 9 H). ^{13}C NMR (150 MHz, $CDCl_3$) 172.9, 171.5, 171.2, 156.6, 144.0, 143.7, 141.5, 128.0, 127.3, 125.3, 120.2, 82.7, 67.7, 62.5, 55.0, 51.7, 47.2, 38.2, 28.2. LRMS: ESI-(+) for $C_{26}H_{30}N_2O_8$ *m/z:* 521.34 $[M+Na^+]^+$.

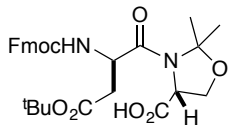


S-11

S-11: Prepared according to an analogous procedure as **S-8**, starting with Fmoc-D-Asp(OtBu)-OH (2.46 mmol, 6.0 mmol). D-Threonine was used for dipeptide formation. Obtained **S-11** as a white solid (2.427 g, 79% yield over 2 steps).

$[\alpha]_D^{22.4} = -23.4$ (c = 1.07, $CHCl_3$)

Fmoc-D-Asp(tBu)-D-Thr-OH **S-11**: ^{1}H NMR (600 MHz, $CDCl_3$) 7.72 (d, *J* = 7.4 Hz, 2 H), 7.49-7.57 (m, 3 H), 7.36 (m, 2 H), 7.25 (m, 2 H), 6.10 (d, *J* = 9.0 Hz, 1 H), 4.68 (m, 1 H), 4.62 (d, *J* = 8.5 Hz, 1 H), 4.43 (m, 2 H), 4.33 (dd, *J* = 10.4, 7.0 Hz, 1 H), 4.19 (app t, *J* = 7.0 Hz, 1 H), 2.79 (dd, *J* = 16.5, 6.5 Hz, 1 H), 2.68 (dd, *J* = 16.4, 4.7 Hz, 1 H), 1.40 (s, 9 H), 1.19 (d, *J* = 6.3 Hz, 3 H). ^{13}C NMR (150 MHz, $CDCl_3$) 173.3, 171.6, 171.2, 156.6, 144.0, 143.7, 141.5, 120.2, 82.4, 68.5, 67.7, 57.7, 51.7, 47.2, 38.1, 28.2, 19.5. LRMS: ESI-(+) for $C_{27}H_{32}N_2O_8$ *m/z*: 535.34 $[M+Na^+]^+$.

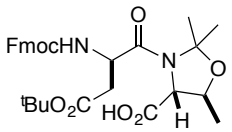


S-12

S-12: Prepared using an analogous procedure for the preparation of **S-9**, starting with **S-10** (2.00 g, 4.01 mmol). Obtained **S-12** as a white solid (1.34 g, 62% yield).

$[\alpha]_D^{22.0} = +50.1$ (c = 0.47, $CHCl_3$)

Fmoc-D-Asp(tBu)-D-Ser($\psi^{Me,Me}$Pro)-OH **S-12**: ^{1}H NMR (600 MHz, DMSO-d_6) Rotameric Mixture 7.89 (d, *J* = 7.5 Hz, 2 H), 7.71 (m, 2 H), 7.60 (d, *J* = 9.1 Hz, 1 H), 7.41 (app t, *J* = 7.5 Hz, 2 H), 7.31 (m, 2 H), 4.63 (m, 1 H), 4.56 (m, 1 H), 4.11-4.32 (m, 5H), 2.74 (dd, *J* = 15.8, 6.4 Hz, 1H), 2.40 (dd, *J* = 15.8, 7.3 Hz), 1.59 (s, 3 H), 1.42 (s, 3 H), 1.37 (m, 9 H). ^{13}C NMR (150 MHz, DMSO-d_6) 171.6, 169.3, 167.4, 155.2, 143.9, 143.7, 140.7, 140.6, 127.6, 127.0, 125.4, 125.3, 120.1, 95.5, 80.2, 67.1, 65.9, 58.7, 50.5, 49.2, 46.5, 38.3, 27.6, 25.2, 22.7. LRMS: ESI-(+)-MS for $C_{29}H_{34}N_2O_8$ *m/z*: 539.34 $[M+H^+]^+$.



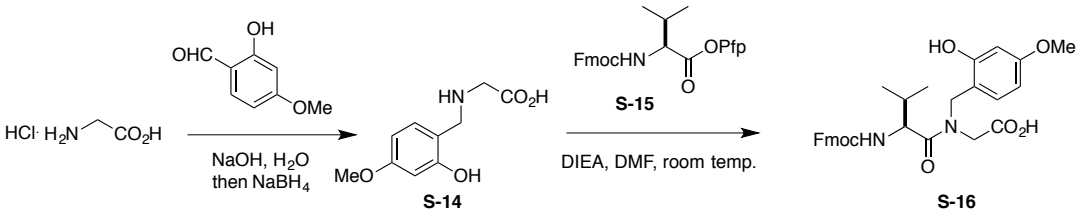
S-13

S13: Synthesized using an analogous procedure for the preparation of **S-9**, starting with **S-11** (1.20 g, 2.34 mmol). Obtained **S-13** as a white solid (908 mg, 70% yield).

$[\alpha]_D^{21.2} = +22.3$ (c = 0.22, $CHCl_3$)

Fmoc-D-Asp(tBu)-Thr($\psi^{Me,Me}$Pro)-OH **S-13**: ^{1}H NMR (600 MHz, DMSO-d_6) Rotameric Mixture 7.90 (d, 7.5 Hz, 2 H), 7.30 (m, 2 H), 7.43 (app t, *J* = 7.5 Hz, 2 H), 7.33 (m, 2 H), 4.48 (m, 1 H), 3.95-4.28 (m, 5 H), 2.65 (dd, *J* = 15.5, 6.1 Hz, 1H), 2.48 (dd, *J* = 15.5, 7.1 Hz, 1 H), 1.54 (s, 3 H), 1.49 (s, 3 H), 1.39 (m, 12 H). ^{13}C NMR (150 MHz, DMSO-d_6) major peaks 169.1, 168.1, 155.7, 155.0, 143.9, 143.7, 142.5, 140.7, 140.6 (2 C), 139.4, 137.4, 128.9, 127.6, 127.3, 127.1, 127.0, 125.4, 125.3, 124.2,121.4, 120.1, 120.0, 109.7, 95.6, 94.0, 80.7, 80.3, 75.0, 72.4, 66.0, 65.7, 51.7, 50.6, 49.4, 46.5, 27.7, 27.6, 27.5, 26.3, 25.9, 24.0, 20.3. LRMS: ESI-(+)-MS for $C_{30}H_{36}N_2O_8$ *m/z*: 575.40 $[M+Na^+]^+$.

Synthesis of Fmoc-Val-(Hmb)Gly-OH dipeptide (S-16):



S-14 was synthesized as adapted from the procedure of Nicolás and co-workers:[50] HCl-Gly-OH (2.00 g, 17.9 mmol) was dissolved in H_2O (32 mL), and the pH was adjusted to pH~9 (pH paper) using 5 M NaOH. To this stirred solution was added 2-hydroxy-4-methoxybenzaldehyde (2.72 g, 17.9 mmol), and the reaction was stirred at room temperature until all solids had fully dissolved (~3 hours), with periodic addition of 5 M NaOH to maintain a pH ~9-10. The resulting fully homogeneous yellow solution was cooled in an ice bath, and $NaBH_4$ (678 mg, 17.9 mmol, dissolved in 15 mL H_2O) was

added slowly, followed by 10 drops of 5 M NaOH. The reaction was further stirred for 1 hour at this temperature, at which point UPLC/MS analysis indicated full reduction of the intermediary imine and the bright yellow color had faded. The reaction was quenched upon the careful addition of 2M HCl until the solution reaches pH~6-7, whereupon a white precipitate formed. The solution was then left overnight in a fridge, and the solids were filtered, washed with H_2O (20 mL), and dried under high vacuum to afford *N*-Hmb-Gly-OH **S-14** (4.00 g, 90% yield) as an off-white solid, which required no additional purification.

S-14: ^{1}H NMR (600 MHz, AcOH-d_4) 7.21 (d, J = 8.4 Hz, 1 H), 6.52 (d, J = 2.4 Hz, 1 H), 6.48 (dd, J = 8.4 Hz, 2.4 Hz, 1 H), 4.30 (s, 2 H), 3.84 (s, 2 H), 3.76 (s, 3H). ^{13}C NMR (150 MHz, AcOH-d_4) 172.6, 164.0, 158.5, 134.3, 110.1, 107.4, 103.3, 56.4, 49.9, 49.0. LRMS: ESI(+)-MS for $C_{10}H_{13}NO_4$ *m/z:* 212.01 $[M+H]^+$.

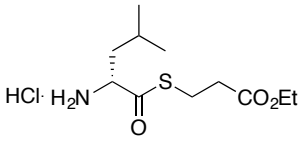
To a 0^{o}C solution of Fmoc-Val-OH (300.0 mg, 0.884 mmol) in EtOAc (8 mL) was added a solution of 2,3,4,5,6-pentafluorophenol (178.0 mg, 0.97 mmol) in EtOAc (2 mL). To this solution was added *N,N'*-dicyclohexylcarbodiimide (199.7 mg, 0.97 mmol), and the reaction was stirred for 4 hours at this temperature. The reaction mixture was then warmed to room temperature and filtered to remove solids. The solids were washed with EtOAc (15 mL), and the filtrate was sequentially washed with sat. aq. $NaHCO_3$ (15 mL) and brine (15 mL). The organic layer was then dried ($MgSO_4$) and concentrated. The resulting crude pentafluorophenol ester (**S-15**) was then dissolved in DMF (4 mL), and *N*-Hmb-Gly-OH **S-14** (280 mg, 1.326 mmol) was added, followed by *N,N*-diisopropylethylamine (0.31 mmol, 1.76 mmol). The reaction was stirred at room temperature for 3.5 hours, and was then diluted with EtOAc (45 mL). The solution was then washed with 0.5 M HCl (15 mL), and the aqueous layer was further extracted with EtOAc (1 x 40 mL). The combined organic layers were washed with brine (2 x 10 mL), dried ($MgSO_4$), and concentrated. Purification by silica gel chromatography (1:1 EtOAc/Hexane +1% AcOH → 2:1 EtOAc/Hexane +1% AcOH) and combination of purest fractions afforded **S-16** as a white solid (229 mg, 49% yield).

Starting from Fmoc-L-Valine: $[\alpha]_D^{22}$ = -23.6^o (c = 0.57, $CHCl_3$)

Starting from Fmoc-D-Valine: $[\alpha]_D^{22} = +23.5^o$ (c = 0.64, $CHCl_3$)

S-16: ^{1}H NMR (600 MHz, $CDCl_3$, 3.6:1 Mixture of amide rotamers) 7.74 (d, *J* = 7.5 Hz, 2 H), 7.53 (m, 2 H), 7.37 (m, 2 H), 7.29 (m, 2 H), 6.95 (d, *J* = 8.2 Hz, 1 H), 6.47 (d, *J* = 2.5 Hz, 1 H), 6.35 (dd, *J* = 8.3, 2.5 Hz, 1 H), 5.56 (d, *J* = 9.5 Hz, 1 H), 4.56 (d, *J* = 15.0 Hz, 1 H), 4.28-4.41 (m, 4 H), 4.15-4.22 (m, 2 H), 3.73 (s, 3 H), 1.98 (m, 1 H), 0.90 (d, *J* = 6.5 Hz, 6 H). ^{13}C NMR (150 MHz, $CDCl_3$) 174.9, 170.6, 162.0, 157.5, 156.8, 143.9, 143.6, 141.5 (2 C), 132.3, 128.0, 127.4, 127.3, 125.4, 125.3, 125.2, 120.2, 113.3, 106.1, 102.8, 67.7, 56.2, 55.5, 48.4, 47.8, 47.2, 32.0, 19.4, 17.7. UPLC/MS: ESI(+)-MS for $C_{30}H_{32}N_2O_7$ *m/z*: 533.32 $[M+H]^+$.

Synthesis of HCl·H-D-Leu-SAlkyl (S-18) and HCl·H-D-Lys(Alloc)-SAlkyl (S-19):

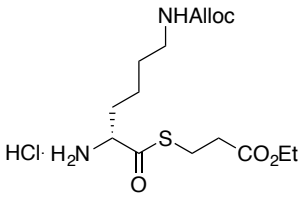


S-18

HCl·H-D-Leu-SAlkyl (**S-18**) was synthesized according to a procedure adapted from Stuhr-Hansen *et al*.[51] To a solution of Boc-D-Leu-OH (1.00 g, 4.32 mmol) in MeCN (13 mL) was added *N,N*-diisopropylethylamine (1.88 mL, 10.8 mmol) followed by HBTU (1.64 g, 4.32 mmol). After 5 minutes, Ethyl 3-mercaptopropionate (1.09 mL, 8.64 mmol) was added. The reaction was stirred until no starting material remained by TLC analysis (5 hours), at which point the mixture was concentrated and purified by silica gel chromatography (100% Hexane→20:1 Hexane/EtOAc→9:1 Hexane/EtOAc) to afford the corresponding thioester as a colorless oil (1.160 g, 77% yield). A portion of the resulting Boc-protected amino acid thioester (1.08 g, 3.11 mmol) was then dissolved in CH_2Cl_2 (3 mL) and treated with 4M HCl in dioxane (6.0 mL, 24.0 mmol) and stirred for 4.5 hours. The reaction was concentrated under reduced pressure, and the residue was dissolved in ca. 5 mL 1:1 MeCN/H_2O (0.1% TFA), shell frozen, and lyophilized to afford HCl-H-D-Leu-SAlkyl **S-18** as a white solid (764 mg, 87% yield).

$[\alpha]_D^{22} = -42.7^o$ (c = 0.38, MeOH)

S-18: ^{1}H NMR (600 MHz, D_2O) 4.33 (dd, *J* = 8.1, 6.3 Hz, 1 H), 4.20 (q, *J* = 7.1 Hz, 2 H), 3.35 (dt, *J* = 13.8, 6.8 Hz, 1 H), 3.25 (dt, *J* = 13.8, 6.8 Hz, 1 H), 2.78 (app t, *J* = 6.8 Hz, 2 H), 1.88 (m, 1 H), 1.77 (m, 2 H), 1.27 (t, *J* = 7.1 Hz, 3 H), 1.00 (d, *J* = 6.4 Hz, 3 H), 0.99 (d, *J* = 6.4 Hz, 3 H). ^{13}C NMR (125 MHz, MeOH-d_4) 198.2, 173.1, 62.1, 59.0, 42.1, 34.9, 25.8, 25.5, 23.0, 22.3, 14.7. LRMS: ESI(+)-MS for $C_{11}H_{21}NO_3S$ *m/z:* 247.75 $[M+H^+]^+$.



S-19

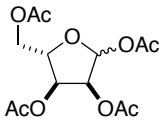
S-19: Synthesized analogously to **S-18**, starting with Boc-D-Lys(Alloc)-OH.[52]

$[\alpha]_D^{22}$ = -41.0° (c = 0.47, MeOH)

S-19: ^{1}H NMR (600 MHz, MeOH-d_4) 5.93 (m, 1 H), 5.28 (dd, *J* = 17.2, 1.4 Hz, 1 H), 5.18 (d, *J* = 10.5 Hz, 1 H), 4.52 (m, 2 H), 4.23 (dd, *J* = 6.8, 5.8 Hz, 1 H) 4.15 (q, *J* = 7.1 Hz, 2 H), 3.26 (m, 2 H), 3.13 (m, 2 H), 2.69 (dt, *J* = 6.7, 1.4 Hz, 2 H), 2.00 (m, 1 H), 1.88 (m, 1 H), 1.56 (m, 2 H), 1.45 (m, 2 H), 1.26 (t, *J* = 7.1 Hz, 3 H). ^{13}C NMR (150 MHz, MeOH-d_4) 197.6, 172.9, 158.9, 134.55, 117.5, 66.3, 62.0, 60.3, 41.1, 34.8, 32.3, 30.5, 25.3, 22.8, 14.5. LRMS: ESI(+)-MS for $C_{15}H_{26}N_2O_5S$ *m/z:* 347.11 $[M+H^+]^+$.

Synthesis of L-GppNHp:

L-GppNHp was synthesized according to literature protocols, starting with L-ribose (Carbosynth). L-Guanosine was synthesized from L-ribose tetraacetate as described by Zou, et al.[18] 5'-L-Guanosine Monophosphate was then converted into the corresponding imidotriphosphate according to the procedure of Yount.[20]



31

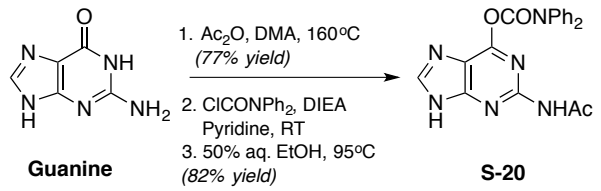
Dowex 8X50W resin (H^+ form, 200-400 mesh, 3.48 g, prewashed with anhydrous MeOH (3 x 10 mL)) was added to a flame-dried 200 mL round bottom flask, and suspended in anhydrous MeOH (70 mL). L-Ribose **29** (3.00 g, 20.0 mmol, purchased from Carbosynth) was added, and the reaction was stirred for 5 hours at room temperature until TLC analysis indicated full consumption of starting material. At this time, the reaction was filtered through a medium porosity frit, the resin washed with additional MeOH (20 mL), and the filtrate was concentrated to yield the crude methyl glycoside which was carried forward without further purification.

The crude residue was dissolved in pyridine (60 mL) and cooled in an ice bath. Acetic anhydride (12.0 mL, 127 mmol) was added, and the reaction stirred for 30 minutes, at which point the ice bath was removed and reaction stirred overnight at room temperature. The reaction mixture was concentrated, taken up in EtOAc (80 mL), and washed successively with water (75 mL), sat. aqueous $NaHCO_3$ (3 x 30 mL, carefully), and Brine (30 mL). The organic layer was dried ($MgSO_4$) and concentrated to yield a residue, which was then coevaporated with dry toluene (3 x 20 mL) and carried forward without further purification.

This crude material (assumed 20.0 mmol) was then dissolved in AcOH (18.0 mL) and Ac_2O (4.5 mL) and cooled in an ice bath. H_2SO_4 (1.26 mL, 23.6 mmol) was added dropwise over 5 minutes, and the reaction stirred 45 minutes, at which point the reaction was warmed to room temperature for 1.5 hours. Additional Ac_2O (3.0 mL) was added, and the reaction stirred 1 hour longer. The reaction mixture was the poured onto 60 mL crushed ice and stirred until all the ice had melted. The mixture was extracted with $CHCl_3$ (3 x 70 mL), and the combined organic layers were washed with sat. aq. $NaHCO_3$ (3 x 40 mL, *carefully*) and then Brine (40 mL). The combined organic layers were then dried ($MgSO_4$) and concentrated. The resulting residue was coevaporated with toluene (3 x 20 mL) then $CHCl_3$ (20 mL) to afford L-Ribose tetraacetate **31** as a colorless oil that solidified upon standing, and which was used without further

purification. (5.590 g, 2.6:1 *dr*, 87% over 3 steps). NMR data and LRMS were consistent with that previously reported.[53]

Synthesis of L-Guanosine and L-Guanosine 5'-monophosphate:



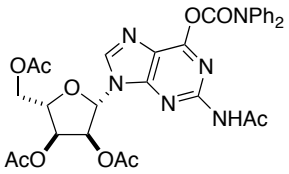
S-20 was synthesized according to the protocol of Zou:[18]

To a suspension of guanine (6.00 g, 39.7 mmol) in anhydrous *N*,*N*-dimethylacetamide (50 mL) was added Ac_2O (10.0 mL, 105.8 mmol). The flask was equipped with a reflux condenser, and stirred in a 160°C oil bath under an argon atmosphere for 6.5 hours. The reaction was cooled to room temperature, and solids were filtered through a Büchner funnel, washed with EtOH (100 mL), and dried under high vacuum to yield bis-acetylated guanine (7.157 g, 77% yield) as a white solid, whose spectroscopic data was consistent with that previously reported.[18]

^{1}H NMR (600 MHz, DMSO-d_6) 8.45 (s, 1 H), 2.81 (s, 3 H), 2.21 (s, 3 H).

A portion of this material (3.06 g, 13.01 mmol) was suspended in pyridine (60 mL), and *N,N*-diisopropylethylamine was added (4.60 mL, 26.4 mmol), followed by solid *N,N*-diphenylcarbamoyl chloride (3.316 g, 14.30 mmol). The reaction was stirred at room temperature for 1 hour, becoming a homogeneous red solution. The reaction was quenched by the addition of H_2O (5.5 mL), and was stirred for a further 10 minutes, at which point solvent was removed under reduced pressure. The residue was coevaporated with toluene (2 x 30 mL) to remove most residual pyridine. To the crude residue was added 50% aqueous EtOH (160 mL), and the mixture was heated under a reflux condenser in a 95°C oil bath for 2 hours. The resulting fine suspension was cooled to room temperature, and solids were filtered and washed with EtOH (3 x 50 mL). The obtained solid was dried overnight in a vacuum desiccator over P_2O_5 to afford pure **S-20** as a white solid (4.166 g, 82% yield). Spectroscopic data matched that previously reported.[18]

S20: ^{1}H NMR (600 MHz, DMSO-d_6) 13.56 (s, 1 H), 10.6 (s, 1 H), 8.44 (s, 1 H), 7.45 (m, 8 H), 7.31 (m, 2 H), 2.15 (s, 3 H).

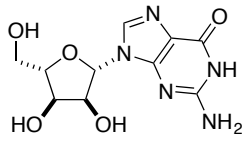


33

Protected guanine derivative **S-20** (1.218 g, 3.14 mmol) was suspended in 1,2-dichloroethane (30.0 mL), and *N,O*-bis(trimethylsilyl)acetamide (1.56 mL, 6.39 mmol) was added. The flask was equipped with a reflux condenser and heated in an 80^{o}C oil bath for approximately 15 minutes, until the solution had become fully homogeneous. After cooling to room temperature, the solvent was removed *in vacuo*, and the residue was redissolved in anhydrous toluene (15.0 mL). To this solution was added TMSOTf (0.79 mL, 2.90 mmol), followed by a solution of L-Ribose tetraacetate **31** (1.199 g, 3.77 mmol, dissolved in 15 mL toluene). The reaction was heated to 80 ^{o}C for 1 hour under argon, then cooled and diluted with 50 mL ethyl acetate. The organic layer was washed with sat. aq. $NaHCO_3$ (30 mL), then brine (50 mL), dried ($MgSO_4$) and concentrated. The residue was purified by silica gel chromatography (1:1 EtOAc/Hexane → 5:1 EtOAc/Hexanes) to yield protected L-guanosine **33** as a colorless foamy solid (1.799 g, 73% yield).

$[\alpha]_D^{21.5} = +4.0^o$ (c = 0.15, $CHCl_3$)

33: ^{1}H NMR (600 MHz, MeOH-d_4) 8.43 (s, 1 H), 7.47 (m, 4 H), 7.39 (m, 4 H), 7.28 (m, 2 H), 6.30 (d, J = 4.4 Hz, 1H), 5.99 (dd, J = 5.8, 4.4 Hz, 1 H), 5.86 (app t, J = 5.4 Hz, 1 H), 4.44 (m, 3 H), 2.27 (s, 3 H), 2.13 (s, 3 H), 2.06 (s, 3 H), 2.00 (s, 3H). ^{13}C NMR (150 MHz, MeOH-d_4) 172.3, 171.8, 171.4, 171.3, 157.2, 155.6, 153.9, 152.1, 145.4, 143.2, 130.4, 128.4, 127.7, 122.2, 88.8, 81.7, 74.7, 72.2, 64.6, 24.7, 20.6, 20.5, 20.3. LRMS: ESI(+)-MS for $C_{31}H_{30}N_6O_{10}$ *m/z:* 647.3 $[M+H^+]^+$, 669.4 $[M+Na^+]^+$

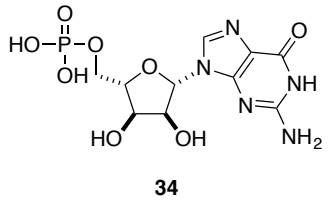


S-21

33 (1.416 g, 2.19 mmol) was dissolved in methanolic ammonia (7N, 20.0 mL), and stirred at room temperature overnight. Addition 7N NH_3 in methanol (10.0 mL) was then added, and the mixture stirred a total of 20 hours. The reaction was directly concentrated, suspended in 80 mL H_2O, and washed with $CHCl_3$ (3 x 50 mL). The combined organic layers were back extracted with 30 mL H_2O to recover emulsified material. The cloudy white aqueous layer was heated in a 100 ^oC oil bath until it became homogeneous, and solids were allowed to precipitate in a refrigerator over 2 days. The precipitated solids were filtered through a Büchner funnel, rinsed with 10 mL cold H_2O, and air-dried to afford L-guanosine (**S-21**) as a white solid (350 mg, 57% yield).

$[\alpha]_D^{22}$ = +72.4^o (*c* = 0.30, 0.1 N NaOH). Lit[54]: $[\alpha]_D^{26}$ = -70^o (*c* = 1.15, 0.1 N NaOH) for enantiomer

S-21: ^{1}H NMR (500 MHz, MeOH-d_4) 7.95 (s, 1 H), 5.83 (d, *J* = 5.9 Hz, 1 H), 4.59 (app t, *J* = 5.5 Hz, 1 H), 4.28 (dd, *J* = 5.1, 3.4 Hz, 1 H), 4.09 (m, 1 H), 3.84 (dd, *J* = 12.3, 2.9 Hz, 1 H), 3.73 (dd, *J* = 12.3, 3.2 Hz, 1 H). ^{13}C NMR (125 MHz, MeOH-d_4) 159.2, 152.3, 152.5, 138.6, 118.4, 90.2, 87.4, 75.5, 72.3, 63.2. LRMS: ESI(+)-MS for $C_{10}H_{13}N_5O_5$ *m/z:* 284.28 $[M+H^+]^+$.



34

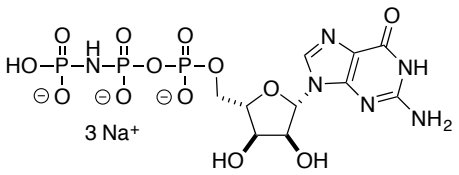
Synthesized according to a previously reported method:[17] **S-24** (177 mg, 0.625 mmol, dried over P_2O_5 in a vacuum desiccator) was suspended in trimethyl phosphate (2.12 mL, dried over 4Å molecular sieves), and heated to 50^{o}C for 15 min. The mixture was

then cooled to 0°C, and H_2O (5.6 uL, 0.31 mmol) was added followed by $POCl_3$ (120 uL, 1.29 mmol) dropwise. The resulting homogeneous light yellow reaction mixture was stirred for 3 hours, and was quenched by the addition of 3 mL H_2O, stirred for 30 minutes, and then adjusted to pH~4-5 with 3M NaOH, and the mixture was shell-frozen and lyophilized. The residue was dissolved in H_2O (10 mL+ 1% TFA), and purified by RP-HPLC. **Conditions:** Microsorb 100-8 C18 column (250 x 21.4 mm), 0-30% MeCN (+0.05% TFA) against H_2O (+0.05% TFA) over 30 minutes, 16 mL/min. The product elutes around 8 minutes. Collected purest fractions are shell-frozen and lyophilized, redissolved in 10 mL H_2O, and a few drops of NH_4OH are added until pH~10. Lyophilization then affords the corresponding ammonium salt **S-25** (133 mg, 58% yield). Spectral data closely matched that reported.[17]

$[\alpha]_D^{21.3}$ = +37.7 (*c* = 0.12, H_2O)

S-25: ^{1}H NMR (600 MHz, D_2O) 8.10 (s, 1 H), 5.93 (d, *J* = 6.1 Hz, 1H), 4.74 (dd, *J* = 5.8, 5.4 Hz, 1 H), 4.47 (dd, *J* = 5.1, 3.5 Hz, 1 H), 4.11 (m, 2 H). ^{13}C NMR (150 MHz, D_2O) 158.9, 153.9, 151.7, 137.5, 116.2, 86.8, 83.9 (d, $J_{C\text{-}P}$ = 8.7 Hz), 73.8, 70.4, 64.4 (d, $J_{C\text{-}P}$ = 4.9 Hz). ^{31}P NMR (200 MHz, D_2O) -2.74 (s, 1 P). LRMS: ESI(+)-MS for $C_{10}H_{14}N_5O_8P$ *m/z:* 364.01 $[M+H^+]^+$; 727.18 $[2M+H^+]^+$.

This nucleotide monophosphate was then converted to its corresponding tributylammonium salt according to the procedure of Yount:[20] **S-25** (133 mg) was added in a minimal volume of cold H_2O to a column of pre-washed Dowex 50WX8 (H^+ form, 2 cm x 2 cm), eluting with cold deionized water. All fractions containing the desired compound (as judged by direct injection mass spectral analysis of individual fractions) were collected and stirred with excess NBu_3 (2.0 mL), and the mixture was concentrated on a rotatory evaporator. This residue was dissolved in 1:1 MeCN/H_2O and lyophilized to yield the corresponding mono-tributylammonium salt as a white solid (169 mg), which was further dried in a vacuum desiccator over solid KOH and $CaSO_4$.



35

Synthesis of ditetrabutylammonium imidodiphosphate[20]

Tetrasodium imidodiphosphate (158.4 mg, 0.6 mmol) was dissolved in 5 mL cold deionized H_2O and eluted on a column of washed Dowex 50WX8 (50-100 mesh, ~30 meq, ~17mL of wet resin), using cold deionized H_2O as the eluent. All acidic fractions (pH paper) were combined in a rapidly stirring solution of excess NBu_3 (1.45 mL) in ~30mL H_2O in an ice-bath. The resulting mixture was then concentrated on a rotatory evaporator with a water bath temperature <30 ^oC. The residual colorless gum was coevaporated with 3 x 15 mL $CHCl_3$, and dissolved in 1.2 mL dry DMF (0.5 M solution, assuming quantitative conversion).

Synthesis of L-GppNHp (35)[20]

L-Guanosine 5'-Monophosphate tributylammonium salt **34** (60.0 mg, 0.109 mmol) was dissolved in anhydrous DMF (0.54 mL) in an oven-dried 1-dram vial, and NBu_3 was added (54 uL, 0.22 mmol, dried over 4Å molecular sieves) followed by diphenyl chlorophosphate (43 uL, 0.123 mmol). The reaction was stirred at room temperature for 2 hours, at which point cold diethyl ether (3 mL) was added, and the reaction flask was kept over a bed of dry ice for 5 minutes. The vial was placed inside a 50 mL plastic centrifuge tube and centrifuged (5 min, 5000 rpm, 10^oC), and the solvent was carefully pipetted out of the vial. The residue was triturated once more with 2 mL ether, centrifuged and decanted to afford a white solid residue. The vial was briefly placed on high vacuum to remove residual ether, and the residue was redissolved in dry DMF (0.60 mL). Separately, ditetrabutylammonium imidodiphosphate (0.5 M solution in DMF, 0.44 mL, 0.22 mmol) was added to a flame-dried 10 mL round bottom flask under argon, and pyridine (1.50 mL) was added. The solution of activated intermediate was then added dropwise via syringe over a 10-minute period, and the reaction was stirred for 2 hours longer (the reaction mixture goes from homogeneous to cloudy over the course of the reaction). The reaction mixture was transferred in two equal portions into two 15 mL centrifuge tubes and precipitated upon the addition of cold diethyl ether

(26 mL total). The suspensions were cooled over a bed of dry ice for 5 minutes, and then centrifuged. The ether was decanted, and the residue triturated with 8 mL cold ether, and decanted again. The resulting white solid was dissolved in 0.1 M triethylammonium bicarbonate (TEAB) buffer (pH ~ 7.5) and purified by ion-pairing RP-HPLC. Conditions: xBridge Prep C18 5µm OBD column (19 x 150 mm), 0-10% MeCN against 0.1 M TEAB (pH ~7.5) over 45 min, 5 mL/min. T_r = 20 min. 1 mL fractions were collected around the eluting peak and analyzed by diluting small aliquots in 4:1 H_2O/MeOH and analyzing by direct injection ESI-MS(-). Fractions containing the desired mass were combined and concentrated on a rotatory evaporator below 30°C, followed by coevaporation of the residue with MeOH (2 x 5 mL).

The resulting triethylammonium salt was converted to the Na^+ form through the following procedure: The above material was dissolved in MeOH (0.8 mL) and transferred via syringe into a 15 mL centrifuge tube. The nucleotide triphosphate was precipitated upon the addition of NaI (1M solution in acetone, 2.5 mL). Additional cold acetone (1.5 mL) was added, the mixture was cooled on a bed of dry ice for 5 minutes, and then centrifuged. The solvent was decanted, and the pellet was further triturated with NaI (1M in acetone) as before, and the mixture was centrifuged as before. The white solid residue was then sequentially triturated and centrifuged with cold acetone (3 x 5 mL), then cold diethyl ether (2 x 5 mL), and finally dried *in vacuo* to yield a white free-flowing solid (14.8 mg, 23% yield). Redissolution and lyophilization from pure H_2O should be avoided, as some hydrolysis of the imidotriphosphate was observed upon relyophilization or standing in water over time (^1H, ^{13}C, and ^{31}P NMR data were consistent with a sample of commercially available GppNHp (Santa Cruz Biotechnology).

^1H NMR (600 MHz, D_2O) 8.12 (s, 1H), 5.92 (d, J = 6.1 Hz, 1 H), 4.81 (app t, J = 5.8 Hz, 1 H), 4.61 (dd, J = 5.0, 3.5 Hz, 1 H), 4.35 (m, 1 H), 4.26 (m, 1 H), 4.20 (m, 1 H). ^{13}C NMR (150 MHz, D_2O) 159.9, 154.6, 151.8, 137.57, 116.4, 86.6, 83.8 (d, $J_{C\text{-}P}$ = 8.6 Hz), 73.4, 70.1, 64.8 (d, $J_{C\text{-}P}$ = 4.3 Hz). ^{31}P NMR (200 MHz, D_2O) -4.20 (d, J = 4.0 Hz, 1 P), -11.65 (br s, 1 P), -13.82 (dt, J = 21.0, 4.8 Hz, 1 P). LRMS ESI(-)-MS for $C_{10}H_{17}N_6O_{13}P_3$ (Free acid) *m/z:* 521.31 $[M\text{-}H^+]^-$.

Part VI. Procedures for the preparation of mant-GppNHp and biological assays

Synthesis of mant-GppNHp:

Syntheses of mant-GppNHp and mant-L-GppNHp were performed as generally described for mant-GTP.[55] GppNHp was dissolved to 20 mM in 500 uL ultrapure water, then the pH was adjusted to ~9.5 with 1 M NaOH. The solution was heated to 38 ^{o}C, and N-Methylisatoic anhydride (6 mg / 34 μmol / 3.4 equiv, technical grade, Aldrich) was added while stirring. The pH was maintained between 9 and 10 by periodic addition of 1 M NaOH. After 3 hours, 1 M HCl was added to lower the pH to ~7 and terminate the reaction. Undissolved N-Methylisatoic anhydride was removed by centrifugation and the product was purified by HPLC using a 30 minute 10-60% MeCN gradient on an Agilent Zorbax semi-preparative C18 column. The expected compound has a m9ss of 655.4 and a λ_{max} at 252 nm with ε=21500 $M^{-1}cm^{-1}$ at pH 7. The observed M+H (ESI+) was 656.1 for mant-GppNHp and 656.1 for mant-L-GppNHp, and observed *m/z* (ESI-) was 654.1 for mant-GppNHp and 654.1 for mant-L-GppNHp.

Ras Refolding and Purification. All steps were performed at 4 ^{o}C, using low-binding plastic consumables and taking care to minimize transfers. Lyophilized Ras proteins were dissolved to 500 uM in refolding buffer (50 mM Tris pH 8.0, 75 mM NaCl, 5 mM $MgCl_2$) containing 5 M GdnHCl, 5 mM DTT, and 100 uM guanine nucleotide, then 10 ul (5 nmol) was rapidly diluted 100-fold with 1000 ul of refolding buffer containing 1 mM DTT and 100 uM guanine nucleotide. After 1 hour incubation, the sample was dialyzed (Pierce 66380) for 3+ hours against 250 ml of refolding buffer containing 1 mM DTT, then concentrated to 25 ul (Pierce 88513) and purified by gel filtration using an 8-minute isocratic run in refolding buffer on an Agilent AdvanceBio SEC 300 Å, 2.7 um column (4.6 x 150 mm) at 1.0 ml/min flow rate at 4 ^{o}C on an Agilent 1200 Series HPLC, collecting the monomer fraction. For CD measurements, the running buffer was 20 mM sodium phosphate pH 7.5, 2 mM $MgCl_2$. Proteins were quantified by absorbance at 280 nm using an extinction coefficient of 19,685 $M^{-1}cm^{-1}$ for nucleotide-bound proteins and 11,920 $M^{-1}cm^{-1}$ for apoproteins. Purification of recombinant KRas(G12V) and loading with GppNHp nucleotides, purification of Raf RBD, and synthesis and purification of RDA peptides was performed as described[24].

The sequence of the RDA44 peptides is
PRRPRCPGDAASIAALHAYWARLWNYLYRVS.

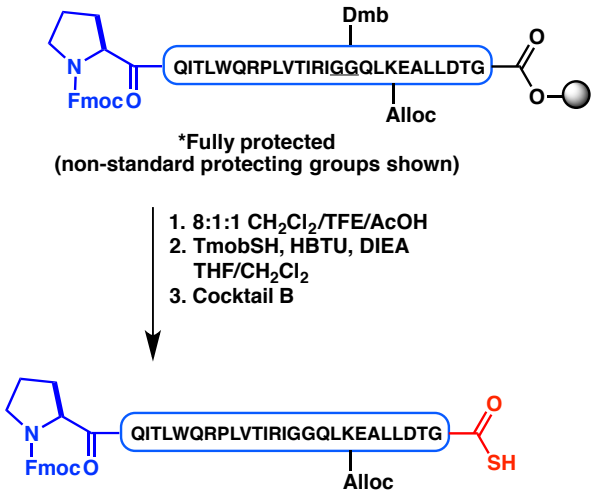
Circular Dichroism (CD) Spectroscopy. CD measurements were carried out on a Jasco J-710 spectrophotometer equipped with a PTC-348W temperature controller. Ras proteins were placed in a 1 mm quartz cuvette at 0.5-1.0 μM in 20 mM sodium phosphate pH 7.5, 2 mM $MgCl_2$, and the CD was scanned from 260 nm to 190 nm at 25 ^{o}C.

Nucleotide Dissociation Assays. mant-GppNHp-bound KRas proteins were prepared by the refolding procedure described above, performing all steps with protection from light. For the assay, the wells of a black 96-well plate (Costar 3915) were seeded with 100 ul of assay buffer (50 mM Tris pH 7.4, 100 mM NaCl, 5 mM $MgCl_2$, 0.1 mg/ml

BSA), containing 2.5 μM Raf or RDA miniprotein, followed by 100 ul of ~0.5 uM purified Ras protein. After 15 minutes incubation to allow binding, nucleotide was added to 10 uM final concentration and the dissociation of mant nucleotide from Ras was monitored by tracking the mant fluorescence (excitation 370 nm, emission 450 nm, cutoff 435 nm) with a SpectraMax M5 (Molecular Devices), recording six reads every 30 seconds.

Part VII: Synthesis of HIV-1 Protease

Synthesis of HIV[1-27] Thioacid (42)



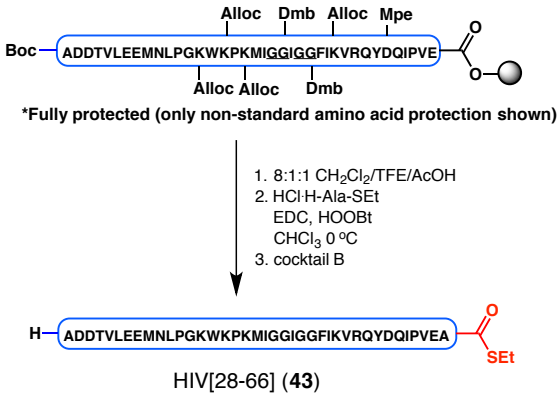
HIV-1 Protease [1-27] (**42**)

Synthesized on 0.075 mmol scale (starting with Fmoc-Gly-OH NovaSyn TGT resin, 0.19 mmol/g, using 5.3 equiv amino acids, 5.3 equiv coupling agents, Deblock: 2% DBU, 2% Piperidine in DMF). Mild cleave using 8:1:1 CH_2Cl_2/TFE/AcOH according to the general procedure afforded the cruded protected peptide.

To a suspension of crude peptide (298 mg, 58 μmol) in THF (8.5 mL) and CH_2Cl_2 (2.0 mL) was added HBTU (110.0 mg, 0.29 mmol), DIEA (140 μL, 0.80 mmol), and 2,4,6-trimethoxybenzylthiol[7] (62.0 mg, 0.29 mmol). The reaction was stirred for 6 hours at room temperature under argon, at which point additional 2,4,6-trimethoxybenzylthiol (37.0 mg, 0.17 mmol) and HBTU (32.0 mg, 0.08 mmol) were added, and the reaction stirred overnight. The mixture was concentrated, and filtered by column chromatography (12 g SiO_2, 0-5% MeOH in CH_2Cl_2, then 10% MeOH in CH_2Cl_2 wash). All fractions containing the desired peptide were concentrated and subjected to cocktail B deprotection (4.5 mL, argon). The peptide was directly precipitated with the addition of cold ether according to the general procedure to afford the crude deprotected peptide. Lyophilization from 1:1 MeCN/H_2O (+0.1% TFA) afforded the crude thioacid as a

white solid (122.5 mg). Due to solubility issues, the residue could not be efficiently purified using RP-HPLC. The thioacid was further purified by reverse-phase column chromatography (12g C2 silica gel, 20 mg portions dissolved in 250 µL TFA). *Conditions:* 15% MeCN→50% MeCN→70% MeCN→90% MeCN against H_2O over 12 minutes. The product elutes around 6.5 min. Combination of the purest fractions and lyophilization afforded thioacid **42** as a white fluffy solid (19.8 mg).
ESI-MS(+) for peptide **42**. Chemical Formula = $C_{156}H_{244}N_{38}O_{41}S$. MW = 3339.95 g/mol. ESI calculated for $[M+2H^+]^{2+}$ *m/z*: 1670.97, found: 1670.99; $[M+3H^+]^{3+}$ *m/z*: 1114.31, found: 1113.91.

Synthesis of HIV[28-66] (43):



Synthesized on a 0.1 mmol scale, starting with Fmoc-Glu(^{t}Bu)-OH NovaSyn TGT Resin. Non-standard amino acids (used at the indicated positions) were Fmoc-Gly-(Dmb)-Gly-OH, Fmoc-Asp(Mpe)-OH, and Fmoc-Lys(Alloc)-OH. Cleavage with 8:1:1 CH_2Cl_2/TFE/AcOH according to the general protocol yielded the desired protected peptide.

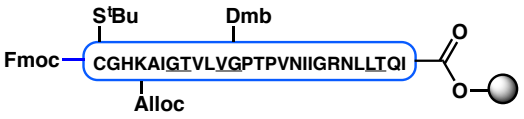
To a suspension of the protected peptide (500 mg, 76.1 µmol) in $CHCl_3$ (7.0 mL) at 0 ^{o}C was added HOOBt (37.2 mg, 0.228 mmol) followed by HCl-H-Ala-Set (34.6 mg, 0.204 mmol) and EDC (40.3 µL, 0.228 mmol). The reaction was stirred at this temperature for 2 hours, and was quenched upon the addition of 6 mL $CHCl_3$ (containing 5% AcOH). The mixture was washed with H_2O (4 mL), which was reextracted once with 5 mL $CHCl_3$ (containing 5% AcOH). The organic layers were

concentrated and subjected to cocktail B (10 mL, 2.5 hr, argon). The peptide was precipitated with cold diethyl ether according to the general protocol, and the pellet was redissolved in 28 mL 1:1 MeCN/H_2O (+0.1% TFA). This solution was allowed to sit at room temperature overnight to allow for any CO_2 adducts with Trp to be removed. Purification of ½ of this material by RP-HPLC afforded the desired peptide **43** as a white solid (31.5 mg, approx. 18% yield).

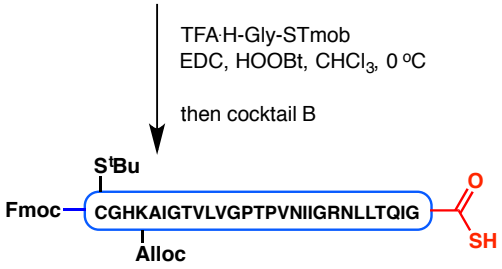
HPLC Conditions: Microsorb 300-5 C4 (250 x 21.4 mm), 30-80% MeCN/H_2O (0.05% TFA) over 30 minutes. *Flow Rate*: 20 mL/min. T_r = 15 minutes.

ESI-MS(+) for peptide **43**: Chemical Formula = $C_{214}H_{332}N_{50}O_{64}S_3$. MW = 4725.48 g/mol. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1576.16, found: 1575.83; $[M+4H^+]^{4+}$ *m/z*: 1182.37, found: 1182.82.

Synthesis of HIV[67-94] Thioacid (44):



*Fully protected (only non-standard amino acid protection shown)



HIV[67-94] (**44**)

Synthesized on a 0.1 mmol scale. Non-standard amino acids used were Fmoc-Lys(Alloc)-OH, Fmoc-Val-(Dmb)Gly-OH, Fmoc-Gly-Thr($\Psi^{Me,Me}$pro)-OH, and Fmoc-Leu-Thr($\Psi^{Me,Me}$pro)-OH. The peptide was cleave from the resin using 8:1:1 CH_2Cl_2/TFE/AcOH according to the general procedure.

Crude peptide (309 mg, 66 µmol) was dissolved in $CHCl_3$ (6.0 mL) and TFE (0.7 mL). The mixture was cooled to 0 °C and HOOBt (32.2 mg, 0.198 mmol) was added followed by TFA·H-Gly-STmob[7] (69.0 mg, 0.18 mmol) and EDC (35 µL, 0.198 mmol).

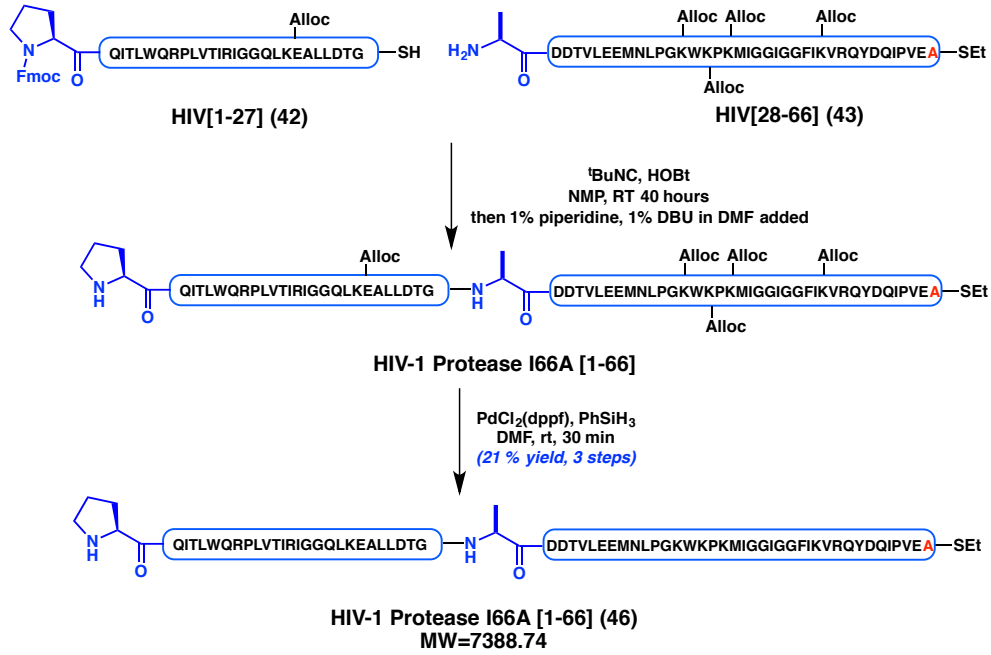
The reaction was stirred for 2.5 hours, at which point some unreacted starting material remained. Additional TFA·H-Gly-STmob (30 mg, 0.078 mmol) and EDC (15 µL, 0.085 mmol) were added and the reaction stirred a further 2 hours. The reaction was quenched upon the addition of 15 mL $CHCl_3$ (+5% AcOH), and the organic layer was washed with H_2O (6 mL) and then concentrated. The peptide was deprotected using cocktail B (8 mL, 2.5 hr, argon) and was precipitated upon addition of diethyl ether according to the general procedure. Due to the insolubility and difficulty in handling of this crude peptide, only a portion was purified by RP-HPLC to afford **44** as a white solid (8.0 mg).
HPLC conditions: C4 Microsorb column, 30-80% MeCN (+0.05% TFA) over 30 min. *Flow rate*: 20 mL/min. T_R = 17 min.
ESI-MS(+) for peptide **44**. Chemical Formula = $C_{149}H_{239}N_{37}O_{38}S_3$. MW = 3252.95 g/mol. ESI calculated for $[M+2H^+]^{2+}$ *m/z*: 1627.48, found: 1627.09; $[M+3H^+]^{3+}$ *m/z*: 1085.31, found: 1085.02; $[M+4H^+]^{4+}$ *m/z*: 814.24, found: 813.97.

Synthesis of HIV[95-115] (45):

HIV[95-115] (45)

Synthesized on a 0.05 mmol scale starting with amide Sieber resin (~0.21 mmol/g loading). Cleavage with cocktail B (5 mL, 5 hours) and precipitation of the peptide with diethyl ether according to the general procedure afforded the crude deprotected peptide. Purification by RP-HPLC afforded **45** as a white solid (11.1 mg, 9% yield).
HPLC Conditions: C8 xBridge column 5 µm OBD, 20-65% MeCN (+0.05% TFA) over 30 min. *Flow rate:* 16 mL/min. T_R = 8.5 min.
ESI-MS(+) for peptide **45**. Chemical Formula = $C_{115}H_{201}N_{41}O_{27}S_2$. MW = 2654.25 g/mol. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 885.75, found: 885.40; $[M+4H^+]^{4+}$ *m/z*: 664.56, found: 664.46.

Synthesis of HIV-1 Protease [1-66] I66A (46):



[Pro^{1}-Gly^{27}] thioacid 42 (3.20 mg, 0.958 umol) and **[Ala^{28}-Ala^{66}] thioester 43** (8.14 mg, 1.72 umol) were directly weighed into a 1.5 mL eppendorf tube with a stirbar. HOBt (1.46 mg, 9.58 umol, dissolved in 90uL anhydrous NMP) was added followed by ^{t}BuNC (0.24 mg, 2.89 umol, dissolved in 90uL anhydrous NMP). The solid fragments were solubilized with stirring, the vial centrifuged for 1 minute, and the headspace of the reaction flushed with argon for 1 minute. The reaction was then stirred at room temperature for 48 hours, at which point a solution containing 1% DBU/1% Piperidine in DMF (300 uL) was added. After 40 minutes, UPLC/MS analysis indicated complete Fmoc removal. The solution was then transferred via micropipette into a 15 mL centrifuge tube and transfer was quantitated with 100uL DMF to wash the reaction vial. Cold ether (10 mL) was added, and the suspension was cooled over dry ice for 5 minutes and then centrifuged. Ether was decanted, and the crude precipitate was washed once more with ether, centrifuged and decanted. The residue was lyophilized from 1:1 MeCN/H_2O (+0.1% TFA).

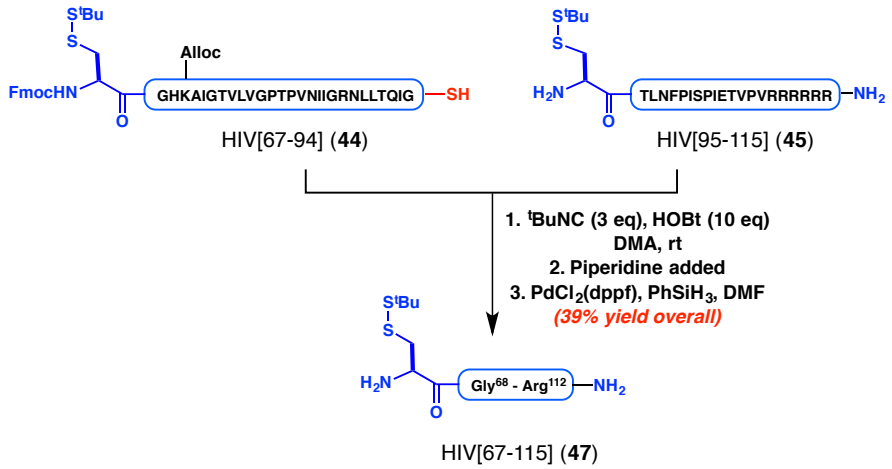
The crude lyophilized material in a 15mL conical centrifuge tube was then dissolved in anhydrous DMF (400 uL, degassed with argon sparge for 30 minutes) and a stirbar was

added. Argon was blown over the headspace of the reaction for 30 seconds, and the vial was centrifuged. $PhSiH_3$ (12.8 uL, 103.8 umol) was added followed by $PdCl_2$(dppf) (0.34 mg, 0.47 umol, freshly prepared as a stock solution in 106 uL degassed DMF). The reaction was stirred for 35 minutes at room temperature (as Alloc deprotection proceeds, the reaction mixture goes from homogeneous to cloudy), at which point UPLC/MS analysis indicated complete deprotection. The reaction was cooled over dry ice for 5 minutes, and cold ether (10 mL) was added to precipitate the crude product. The suspension was centrifuged and ether was decanted. The crude solid was again washed with ether (5 mL), centrifuged, and ether was decanted. The crude peptide was redissolved in 2mL 1:1 MeCN/H_2O (+0.1% TFA) and purified by RP-HPLC. Pure fractions were combined and lyophilized to afford **46** as a white fluffy solid (1.55 mg, 21% overall yield).

HPLC Conditions*:* C8 xBridge column OBD 5 μm, 20-60% MeCN (0.05% TFA) against H_2O (0.05% TFA) over 30 minutes. Flow rate: 16mL/min. T_r = 15 min.

ESI-MS(+) for peptide **46**: Chemical Formula = $C_{335}H_{544}N_{88}O_{93}S_3$. MW = 7388.74 g/mol. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1848.18, found: 1847.93; $[M+5H^+]^{5+}$ *m/z*: 1478.75, found: 1478.64; $[M+6H^+]^{6+}$ *m/z*: 1232.45, found: 1232.34; $[M+7H^+]^{7+}$ *m/z*: 1056.53, found: 1056.43; $[M+8H^+]^{8+}$ *m/z*: 924.59, found: 924.43; $[M+9H^+]^{9+}$ *m/z*: 821.97, found: 821.55; $[M+10H^+]^{10+}$ *m/z*: 739.87, found: 739.37.

Synthesis of HIV-1 Protease [67-115] (47):



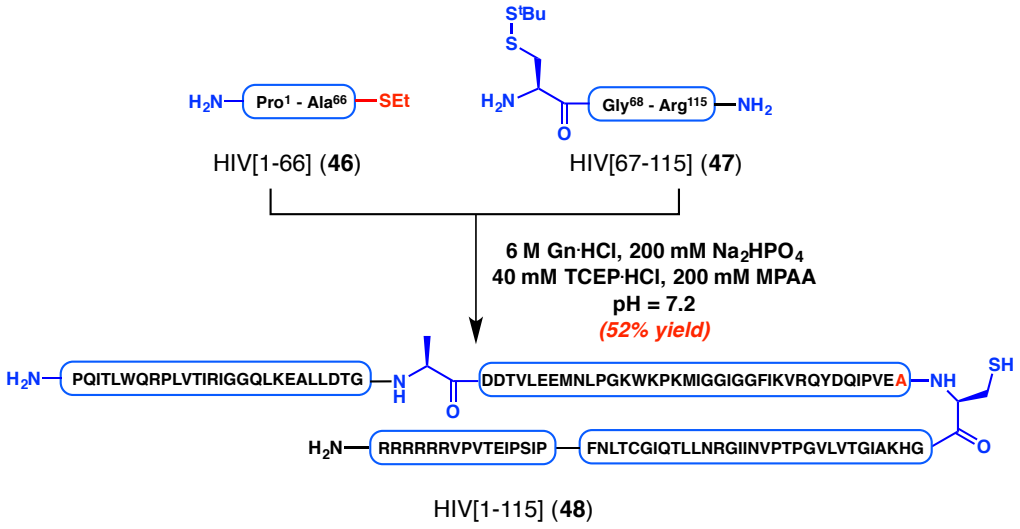
Thioacid 44 (1.78 mg, 0.547 umol) and **45** (2.17 mg, 0.82 umol) were weighed directly into a 1.5 mL eppendorf tube. HOBt (0.84 mg, 5.47 umol, prepared as a stock solution in 55uL anhydrous DMA) was added followed by ^{t}BuNC (0.13 mg, 1.64 umol, prepared as a stock solution in 55 uL anhydrous DMA). Argon was blown over the headspace of the reaction for 30 seconds, and the reaction was sealed and stirred for 48 hours at room temperature. At this time, piperidine (15 uL) was added, and the reaction stirred a further 20 minutes, at which point UPLC/MS analysis showed complete Fmoc deprotection. Cold ether (1 mL) was added to precipitate the crude peptides, the reaction vessel was centrifuged and ether was decanted. The precipitate was washed once more with ether (1 mL), centrifuged and decanted as before. The pellet was dissolved in 0.7 mL 1:1 MeCN/H_2O (+0.1% TFA) and lyophilized.

The crude lyophilized peptide was then suspended in anhydrous DMF (150 uL, degassed with argon sparge for 30 minutes). $PhSiH_3$ (1.4 uL, 10.9 umol) was added followed by $PdCl_2$(dppf) (0.08 mg, 0.11 umol, freshly prepared as a stock solution in 20 uL degassed DMF). The mixture was stirred for 30 minutes, at which point cold ether (1 mL) was added to precipitate the crude peptide. The suspension was centrifuged, and the ether was decanted. The residue was then redissolved in 0.7 mL 1:1 MeCN/H_2O (+0.1% TFA) and purified by RP-HPLC. Combined pure fractions were lyophilized to afford **47** as a white solid (1.20 mg, 39% yield).

HPLC Conditions: C8 X-Bridge Column 5 µm OBD, 20-65% MeCN (+0.05% TFA) against H_2O (+0.05% TFA). *Flow rate:* 15mL/min. T_r = 16 min.

ESI-MS(+) for peptide **47**: Chemical Formula = $C_{245}H_{424}N_{78}O_{61}S_4$. MW = 5566.81 g/mol. ESI calculated for $[M+3TFA+3H^+]^{3+}$ *m/z*: 1970.62, found: 1970.39; $[M+2TFA+4H^+]^{4+}$ *m/z*: 1449.71, found: 1449.40; $[M+1TFA+4H^+]^{4+}$ *m/z*: 1421.21, found: 1421.08; $[M+4H^+]^{4+}$ *m/z*: 1392.70, found: 1392.56; $[M+5H^+]^{5+}$ *m/z*: 1114.36, found: 1114.33; $[M+6H^+]^{6+}$ *m/z*: 928.80, found: 928.92; $[M+7H^+]^{7+}$ *m/z*: 796.26, found: 796.42.

Native Chemical Ligation for HIV-1 Protease I66A [1-115] (48)



HIV[67-115] **47** (1.20 mg, 0.21 µmol) was dissolved in 100 µL NCL buffer (6M Gn·HCl, 200 mM Na_2HPO_4, 200 mM MPAA, 40 mM TCEP·HCl, pH = 7.2, degassed 30 minutes with argon sparge and sonication) in a 1.5 mL eppendorf tube. The tube was vortexed and centrifuged for 1 minute, and the solution transferred via micropipette into a 1.5 mL eppendorf tube containing HIV[1-66] **46** (1.55 mg, 0.21 umol). Transfer was quantitated with 20 uL NCL buffer to wash the initial vial, with vortexing and centrifugation as before. Additional TCEP (4 uL 0.5M aqueous solution, 2 umol, pH = 7.0) was added, argon was blown over the headspace of the reaction for 30 seconds, and the reaction was stirred at room temperature for 17 hours. The reaction was quenched by the addition of TCEP (6 µL, 0.5M aqueous solution, 3 µmol, pH = 7.0) and 500 µL 1:1 MeCN/H_2O (+0.1% TFA). The solution was then purified by RP-HPLC. Pure

fractions were combined and lyophilized to afford **48** as a white solid (1.38 mg, 52% yield).

HPLC Conditions: C8 X-Bridge column 5 μm OBD, 20-55% MeCN (+0.05% TFA) against H_2O (+0.05% TFA) over 30 min. Flow rate: 15 mL/min. T_r = 17.4 min.

ESI-MS(+) for peptide **48**: Chemical Formula = $C_{570}H_{946}N_{166}O_{154}S_4$. MW = 12717.09 g/mol. ESI calculated for $[M+7H^+]^{7+}$ *m/z*: 1817.72, found: 1817.48; $[M+8H^+]^{8+}$ *m/z*: 1590.65, found: 1590.47; $[M+9H^+]^{9+}$ *m/z*: 1414.01, found: 1413.88; $[M+10H^+]^{10+}$ *m/z*: 1272.70, found: 1272.67; $[M+11H^+]^{11+}$ *m/z*: 1157.10, found: 1157.11; $[M+12H^+]^{12+}$ *m/z*: 1060.76, found: 1060.75; $[M+13H^+]^{13+}$ *m/z*: 979.24, found: 979.30; $[M+14H^+]^{14+}$ *m/z*: 909.36, found: 909.39; $[M+15H^+]^{15+}$ *m/z*: 848.80, found: 848.94.

Folding of HIV-1 Protease I66A:

HIV[1-115] **48** (1.38 mg, 0.11 μmol) was dissolved in denaturing buffer (1.5 mL, 6M Gn·HCl, 200 mM Na_2HPO_4, 40 mM TCEP·HCl, pH ~ 7.4). A portion of this solution (100 μL) was diluted with 100 μL folding buffer (100 mM NaOAc, containing 10% glycerol, pH = 5.5). This solution was loaded into a 2 mL Slide-A-Lyzer® MINI dialysis unit (3500 MW cutoff). The solution was dialyzed in a 50 mL centrifuge tube against 50 mL folding buffer submerged in an ice water bath for 2 hours. Buffer was replaced and the sample was dialyzed for a further 1.5 hours. The sample was further dialyzed at this temperature against 50 mL diluted folding buffer (25 mM NaOAc, 10% glycerol, pH = 5.5) for one hour, and the buffer was replaced once more and the sample was dialyzed 1 hour longer. The resulting sample was then removed, centrifuged, and allowed to sit on ice for a further 2 hours. At this point, only a trace amount of unfolded starting material remained by UPLC/MS analysis, with the major peak corresponding to the 99-residue folded protein (**49**).

ESI-MS(+) for peptide **49**: Chemical Formula = $C_{485}H_{793}N_{131}O_{135}S_4$. MW = 10747.70 g/mol. ESI calculated for $[M+6H^+]^{6+}$ *m/z*: 1792.27, found: 1792.28; $[M+7H^+]^{7+}$ *m/z*: 1536.39, found: 1536.21; $[M+8H^+]^{8+}$ *m/z*: 1344.46, found: 1344.38; $[M+9H^+]^{9+}$ *m/z*: 1195.19, found: 1195.11; $[M+10H^+]^{10+}$ *m/z*: 1075.77, found: 1075.64; $[M+11H^+]^{11+}$ *m/z*:

978.06, found: 978.08; $[M+12H^+]^{12+}$ *m/z*: 896.64, found: 896.66; $[M+13H^+]^{13+}$ *m/z*: 827.75, found: 828.15.

REFERENCES

1. Bar-Sagi, D. *Mol. Cell. Biol.* **2001**, *21*, 1441-1443.
2. Nussinov, R.; Muratcioglu, S.; Tsai, C. J.; Jang, H.; Gursoy, A.; Keskin, O. *Mol. Cancer Res.* **2015**, *13*, 1265-1273.
3. Abraham, S. J.; Nolet, R. P.; Calvert, R. J.; Anderson, L. M.; Gaponeko, V. *Biochem.* **2009**, *48*, 7575-7583.
4. Chavan, T. S.; Jang, H.; Khavrutskii, L.; Abraham, S. J.; Banerjee, A.; Freed, B. C.; Johannessen, L.; Tarasov, S. G.; Gaponenko, V.; Nussinov, R.; Tarasova, N. I. *Biophys. J.* **2015**, *109*, 2602-2613.
5. Lu, S.; Banerjee, A.; Jang, H.; Zhang, J.; Gaponeko, V.; Nussinov, R. *J. Biol. Chem.* **2015**, *290*, 28887-28900.
6. Wan, Q.; Danishefsky, S. J. *Angew. Chem. Int. Ed. Eng.* **2007**, *46*, 9248-9252.
7. Roberts, A. G.; Johnston, E. V.; Shieh, J. H.; Sondey, J. P.; Hendrickson, R. C.; Moore, M. A.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2015**, *137*, 13167-13175.
8. Bang, D.; Pentelute, B. L.; Kent, S. B. H. *Angew. Chem. Int. Ed. Eng.* **2006**, *45*, 3985-3988.
9. Bang, D.; Kent, S. B. H. *Angew. Chem. Int. Ed. Eng.* **2004**, *43*, 2534-2538.
10. Sakakibara, S. *Biopolymers* **1995**, *37*, 17-28.
11. DeLoskey, R. J.; Van Dyk, D. E.; Van Aken, T. E.; Campbell-Burk, S. *Arch. Biochem. Biophys.* **1994**, *311*, 72-78
12. Becker, C. F. W.; Hunter, C. L.; Seidel, R.; Kent, S. B. H.; Goody, R. S.; Engelhard, M. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5075-5080.

13. Johnson, T.; Quibell, M.; Owen, D.; Sheppard, R. C. *J. Chem. Soc., Chem. Commun.* **1993**, 369-372.

14. Johnston, T.; Quibell, M. *Tetrahedron Lett.* **1994**, *35*, 463-466.

15. Nicolás, E.; Pujades, M.; Bacardit, J.; Giralt, E.; Albericio, F. *Tetrahedron Lett*. **1997**, *38*, 2317-2320.

16. Zheng, J.-S.; Yu, M.; Qi, Y.-K.; Tang, S.; Shen, F.; Wang, Z.-P.; Xiao, L.; Zhang, L.; Tian, C.-L.; Liu, L. *J. Am. Chem. Soc.* **2014**, *136*, 3695-3704.

17. Mehta, A. P.; Hanes, J. W.; Abdelwahed, S. H.; Himey, D. G.; Hänzelmann, P.; Begley, T. O. *Biochem.* **2013**, *52*, 1134-1136.

18. Zou, R.; Robins, M. J. *Can. J. Chem.* **1987**, *65*, 1436-1437.

19. Yoshikawa, M.; Kato, T.; Takenishi, T. *Tetrahedron Lett*. **1965**, *50*, 5065-5068.

20. Yount, R. G. *Meth. Enzymol.* **1974**, *38*, 420-427.

21. Milić, J.; Seidel, R.; Becker, C. F. W.; Goody, R. S.; Engelhard, M. *Pept. Sci.* **2007**, *90*, 399-405.

22. Gulland, J. M.; Story, L. F. *J. Chem. Soc.* **1938**, 692-694.

23. John, J.; Sohmen, R.; Feuerstein, J.; Linke, R.; Wittinghofer, A.; Goody, R. S. *Biochem.* **1990**, *29*, 6058-6065.

24. McGee, J. H.; Shim, S.; Lee, S.-J.; Swanson, P. K.; Jiang, Y.; Durney, M. A.; Verdine, G. L. *(Submitted)* **2016**.

25. Brik, A.; Wong, C.-H. *Org. Biomol. Chem.* **2003**, *1*, 5-14.

26. De Clercq, E. *Nat. Rev. Drug Disc.* **2007**, *6*, 1001-1018.

27. Schneider, J.; Kent, S. B. H. *Cell* **1988**, *54*, 363-368.

28. Schnölzer, M.; Kent, S. B. *Science* **1992**, *256*, 221-225.

29. Milton, R. C.; Milton, S. C.; Kent, S. B. *Science* **1992**, *256*, 1445-1448.

30. Baca, M.; Muir, T. W.; Schnölzer, M.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 1881-1887.

31. Torbeev, V. Y.; Kent, S. B. H. *Angew. Chem. Int. Ed. Eng.* **2007**, *46*, 1667-1670.

32. Johnson, E. C.; Malito, E.; Shen, Y.; Rich, D.; Tang, W. J.; Kent, S. B. *J. Am. Chem. Soc.* **2007**, *129*, 11480-11490.

33. Liu, C.-F.; Rao, C.; Tam, J. P. *J. Am. Chem. Soc.* **1996**, *188*, 307-312.

34. Nutt, R. F.; Brady, S. F.; Darke, P. L.; Ciccarone, T. M.; Colton, C. D.; Nutt. E. M.; Rodkey, J. A.; Bennett, C. D.; Waxman, L. H.; Sigal, I. S.; Anderson, P. S.; Veber, D. F. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 7129-7133.

35. Qi, Y.-K.; Chang, H.-N.; Pan, K.-M.; Tian, C.-L.; Zheng, J.- S. *Chem. Commun.* **2015**, *51*, 14632-14635.

36. Strickler, J. E.; Gorniak, J.; Dayton, B.; Meek, T.; Moore, M.; Magaard, V.; Malinowski, J.; Debouck, C. *Proteins: Structure, Function, and Genetics* **1989**, *6*, 139-154.

37. Fukuyama, T.; Lin, S.-C.; Li, L. *J. Am. Chem. Soc.* **1990**, *112*, 7050-7051.

38. Katayama, H.; Hojo, H.; Ohira, T.; Nakahara, Y. *Tetrahedron Lett.* **2008**, *49*, 5492-5494.

39. Lapeyre, M.; Leprince, J.; Massonneau, M.; Oulyadi, H.; Renard, P.-Y.; Romieu, A.; Turcatti, G.; Vaudry, H. *Chem. Eur. J.* **2006**, *12*, 3655-3671.

40. Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705-6710.

41. Gude, M.; Ryf, J.; White, P. D. *Lett. Pept. Sci.* **2002**, *9*, 203-206.

42. Creech, G. S.; Paresi, C.; Li, Y.-M.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 2891-2896.

43. Aussedat, B.; Fasching, B.; Johnston, E.; Sane, N.; Nagorny, P.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2012**, *134*, 3532-3541.

44. Zheng, J.–S.; Tang, S.; Qi, Y.–K.; Wang, Z. –P.; Liu, L. *Nature Protocols* **2013**, *8*, 2483-2495.

45. Huang, Y.–C.; Chen, C. –C.; Li, S.–J.; Gao, S.; Shi, J.; Li, Y. –M. *Tetrahedron,* **2014**, *70*, 2951-2955.

46. Nakabayashi, S.; Warren, C. D.; Jeanloz, R. W. *Carbohydrate Res.* **1988**, *174*, 279-289

47. Karlström, A.; Undén, A. *Tetrahedron Lett.* **1996**, *37*, 4243-4246.

48. Procedure for dipeptide synthesis was adapted from a previously described method: Meneses, C.; Nicoll, S. L.; Trembleau, L. *J. Org. Chem.* **2010**, *75*, 564-569

49. Dipeptides were converted to the corresponding pseudoproline dipeptides according to a previously reported method: Wöhr, T.; Wahl, F.; Nefzi, A.; Rohwedder, B.; Sato, T.; Sun, X.; Mutter, M. *J. Am. Chem. Soc.* **1996**, *118*, 9218-9227.

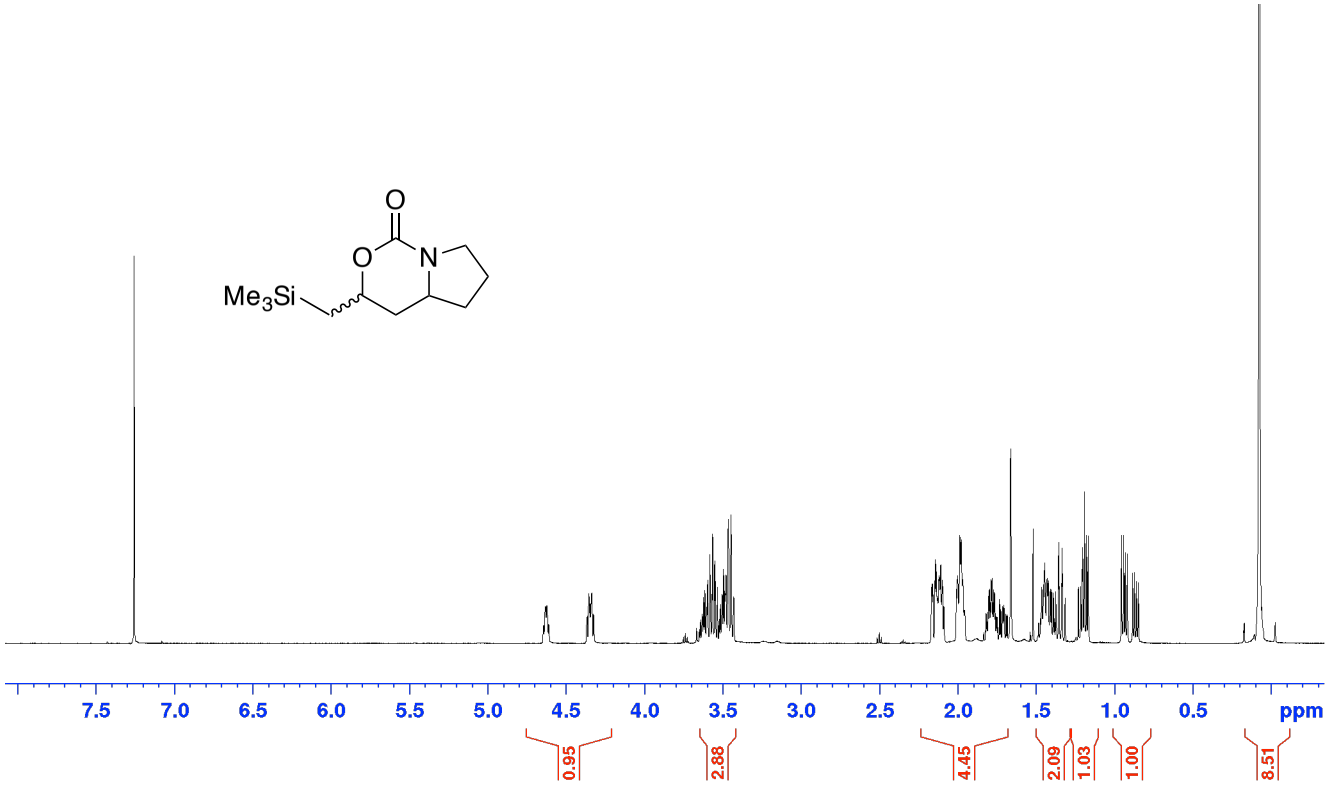
50. Nicolás, E.; Pujades, M.; Bacardit, J.; Giralt, E.; Albericio, F. *Tetrahedron Lett.* **1997**, *38* 2317-2320.

51. Stuhr-Hansen, N.; Wilbek, T. S.; Stømgaard, K. *Eur. J. Org. Chem.* **2013**, 5290-5294.

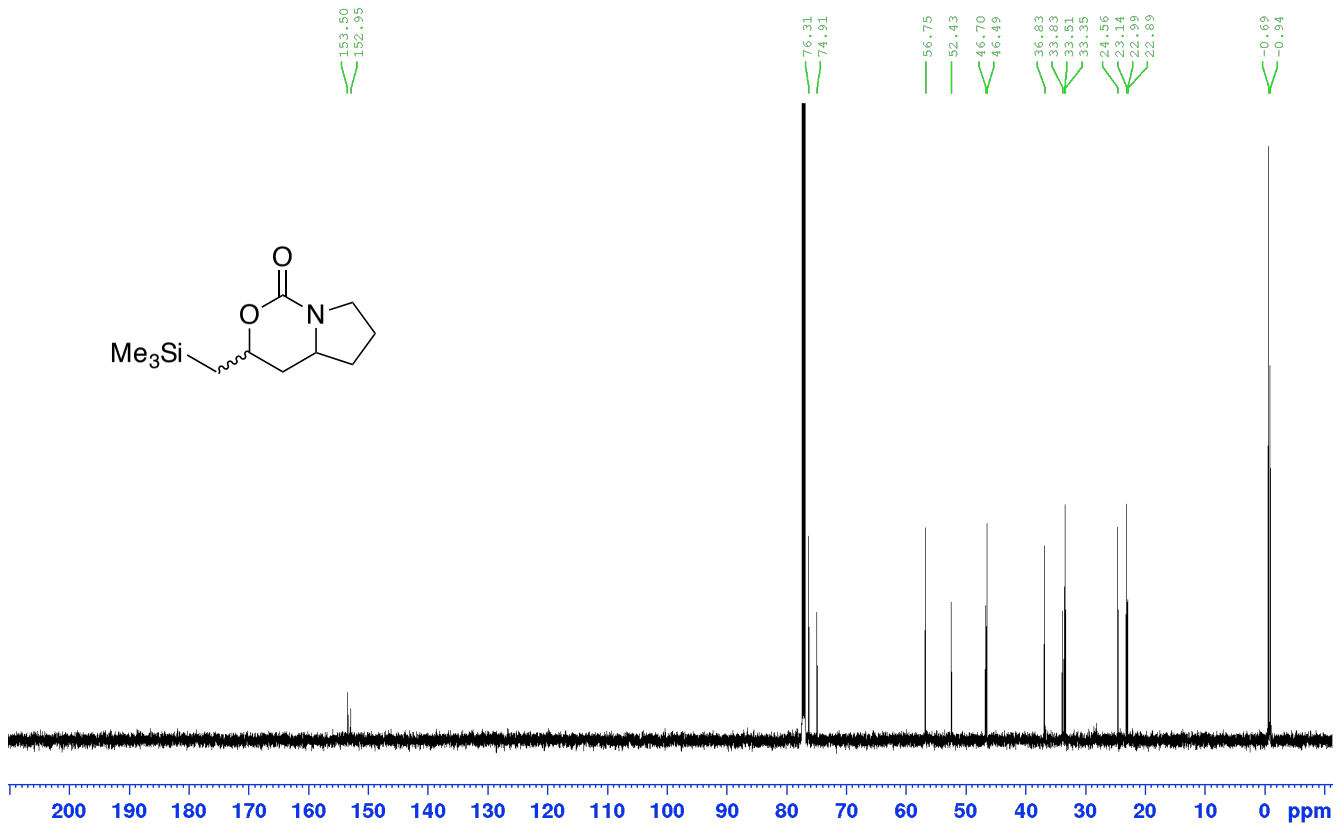
52. Synthesized from Boc-D-Lys-OH according to the previously reported procedure: Ai, H. –W.; Lee, J. W.; Schultz, P. G. *Chem. Commun.* **2010**, *46*, 5506-5508.

53. Patching, S. G.; Baldwin, S. A.; Baldwin, A. D.; Young, J. D.; Gallagher, M. P.; Henderson, P. J. F.; Herbert, R. B. *Org. Biomol. Chem.* **2005**, *3*, 462-470.
54. Davoll, J.; Lowy, B. A. *J. Am. Chem. Soc.* **1951**, *73*, 1650-1655
55. Hiratsuka, T. *Biochim. Biophys. Acta* **1983**, *742*, 496-508.

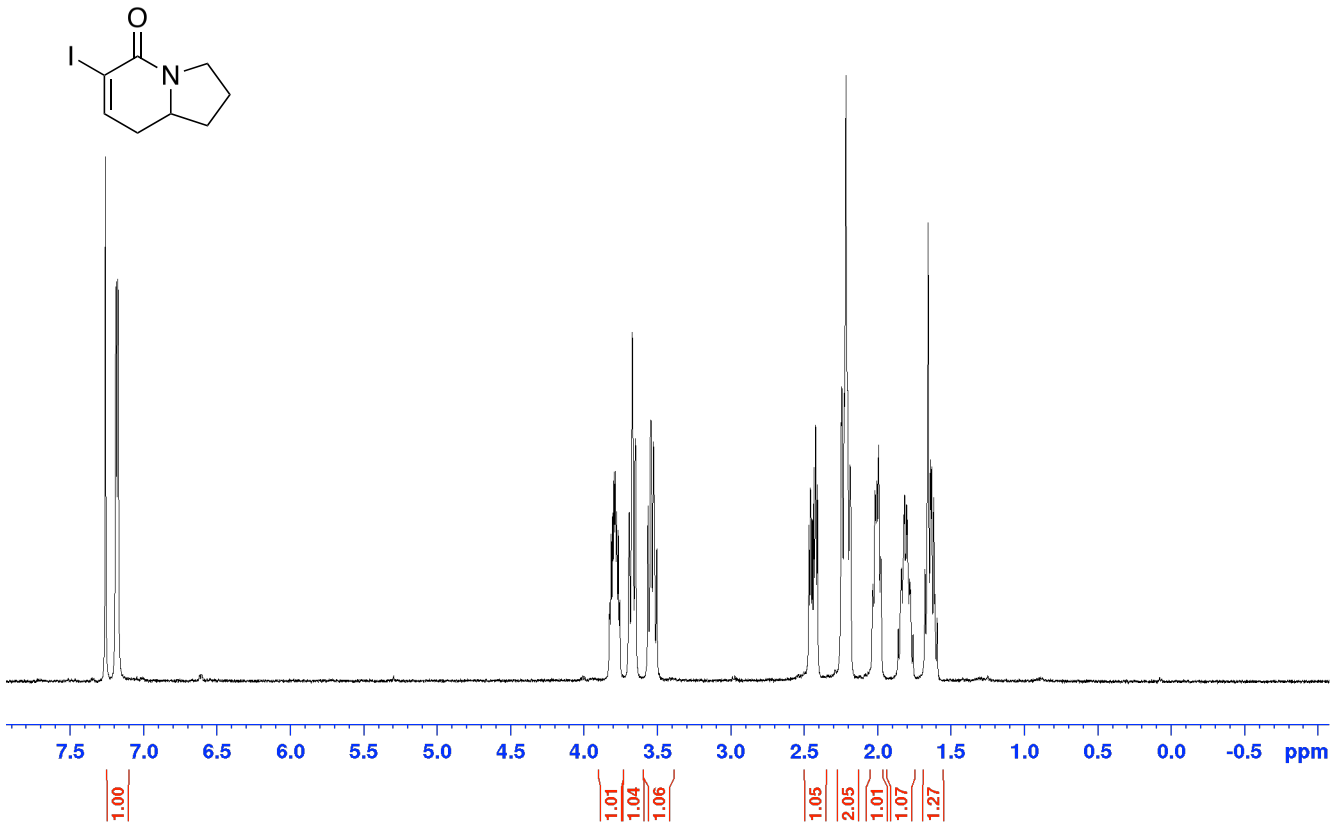
Appendix I :
Selected NMR Spectra for Chapter 2



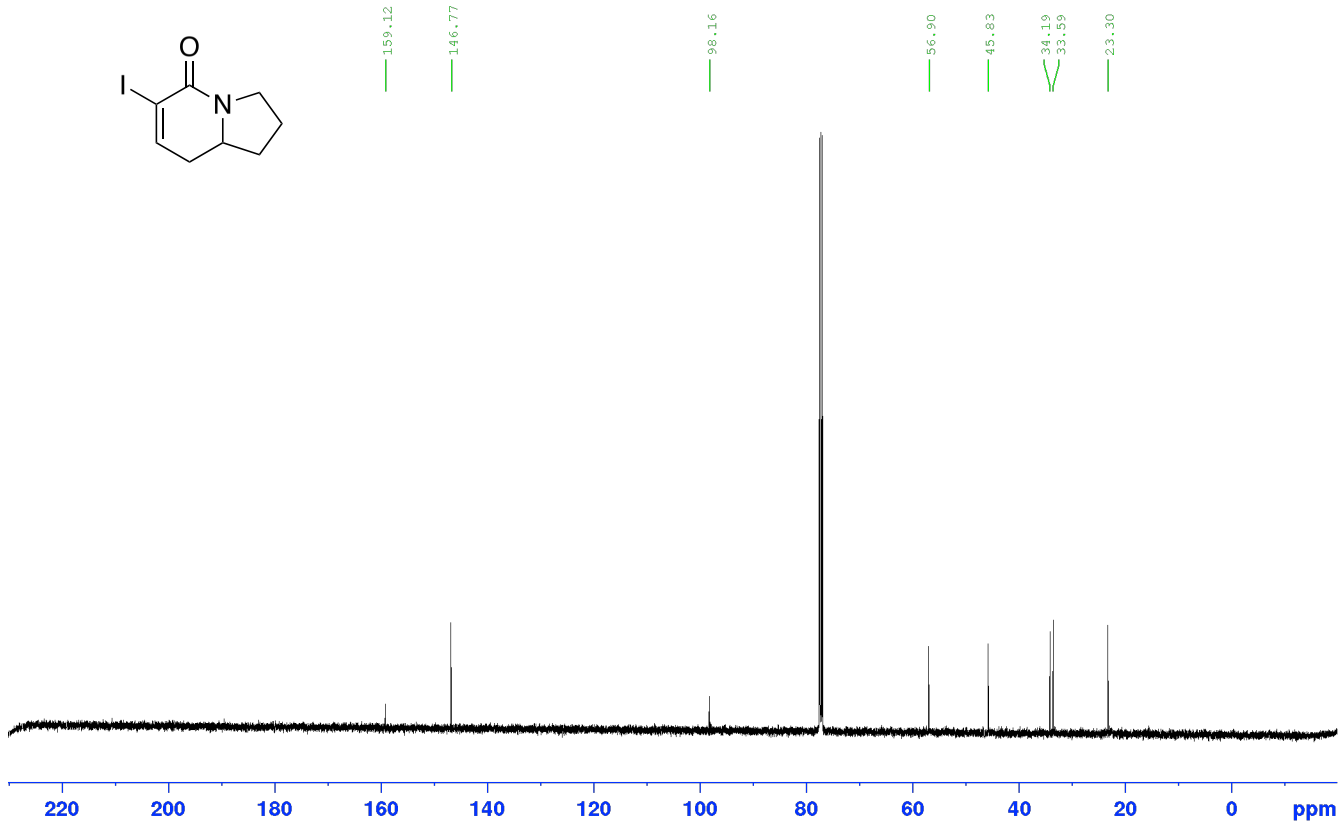
^{1}H NMR of compound **7** (600 MHz, $CDCl_3$)



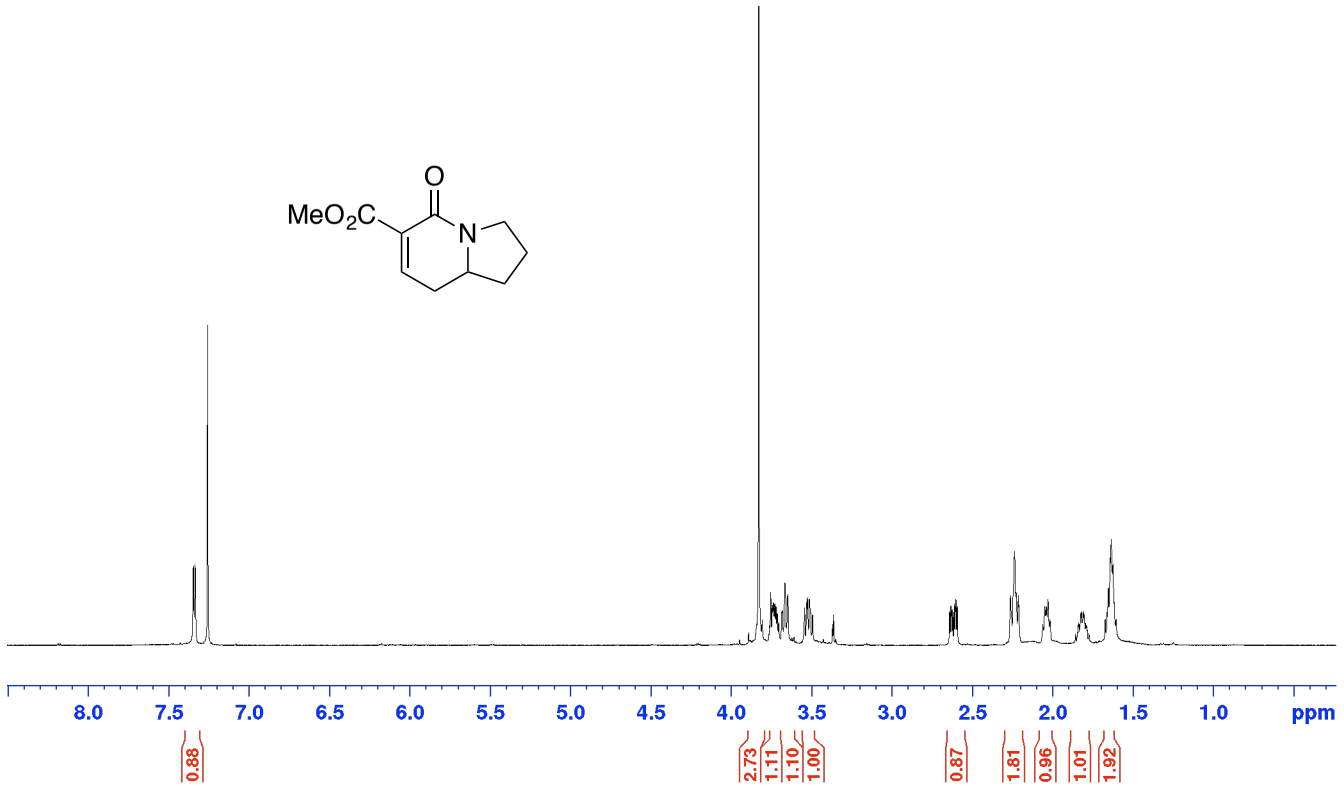
^{13}C NMR of compound **7** (150 MHz, $CDCl_3$)



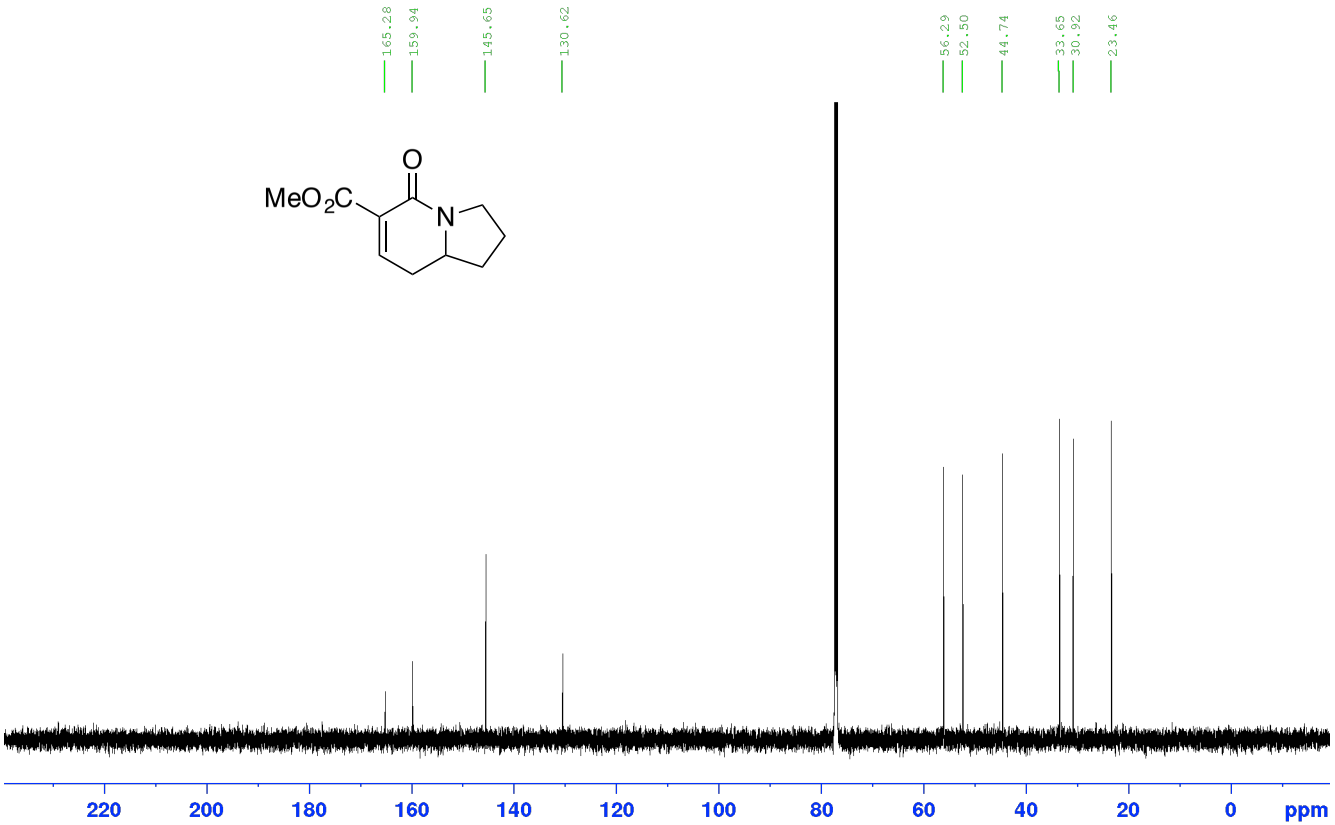
^{1}H NMR of compound **11** (500 MHz, $CDCl_3$)



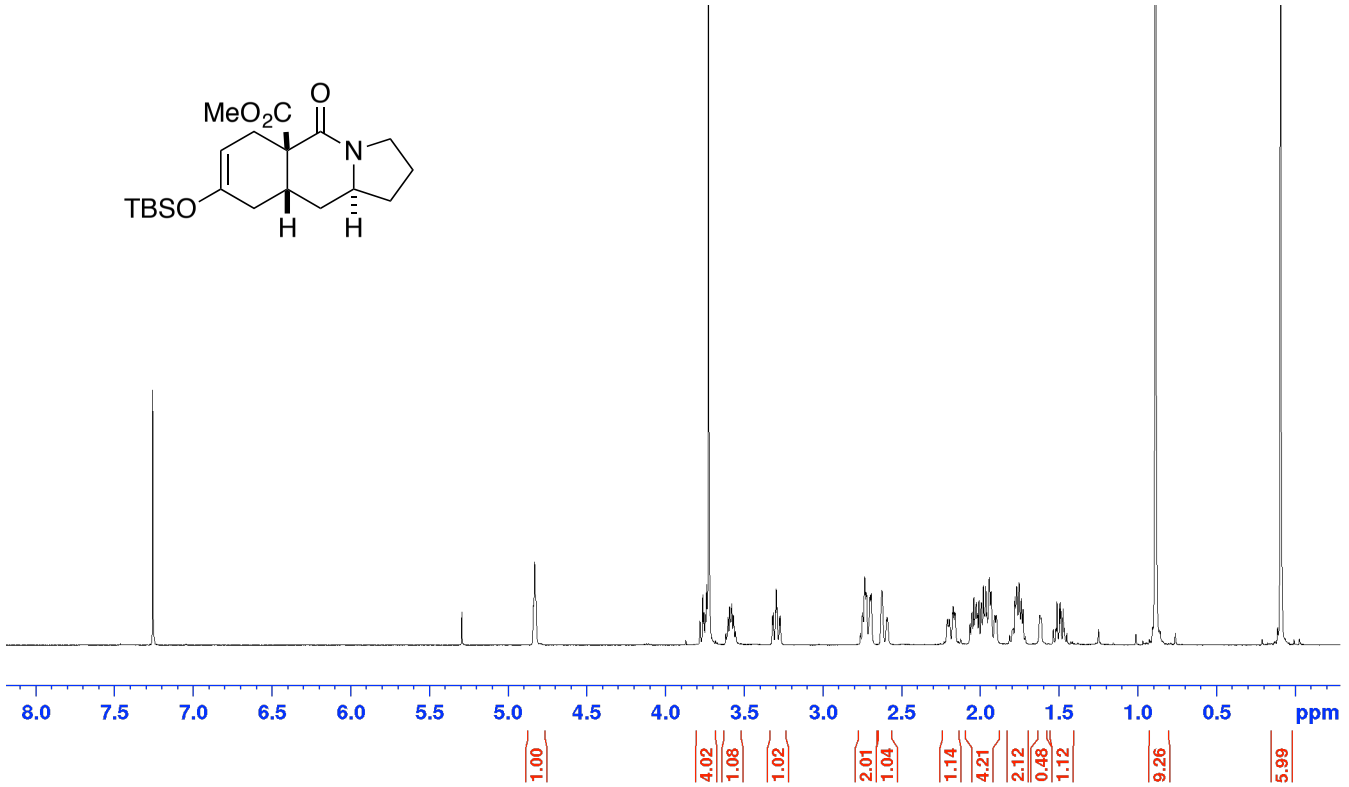
^{13}C NMR of compound **11** (125 MHz, $CDCl_3$)



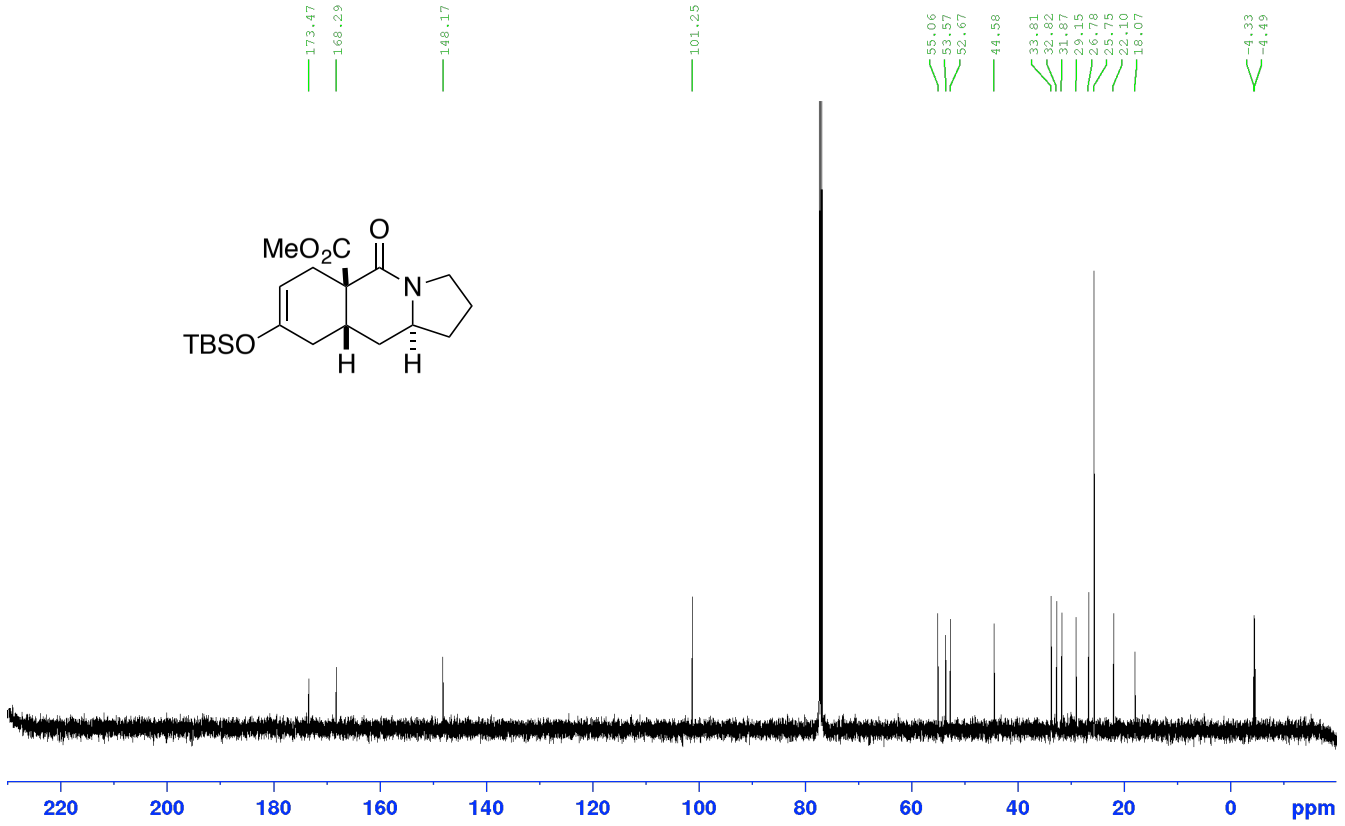
^{1}H NMR of compound **12** (600 MHz, $CDCl_3$)



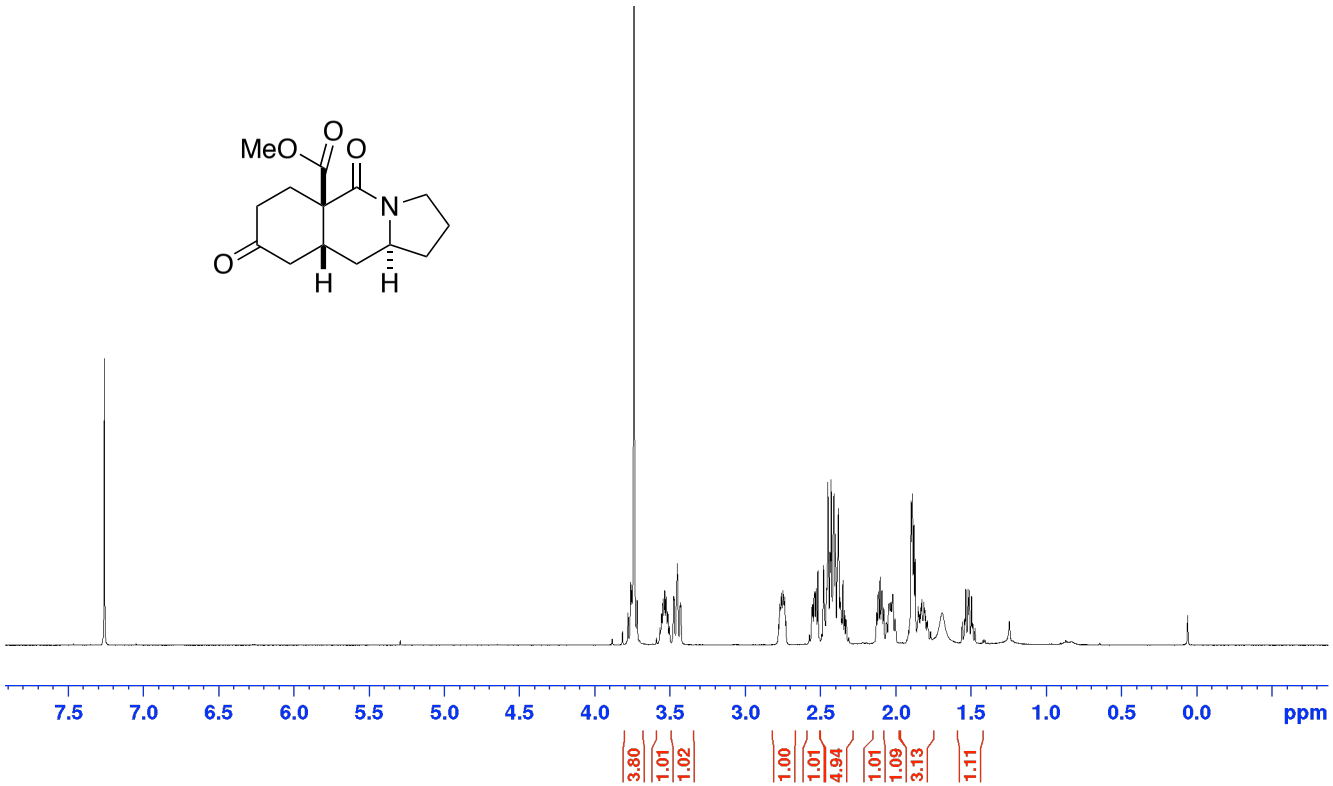
^{13}C NMR of compound **12** (150 MHz, $CDCl_3$)



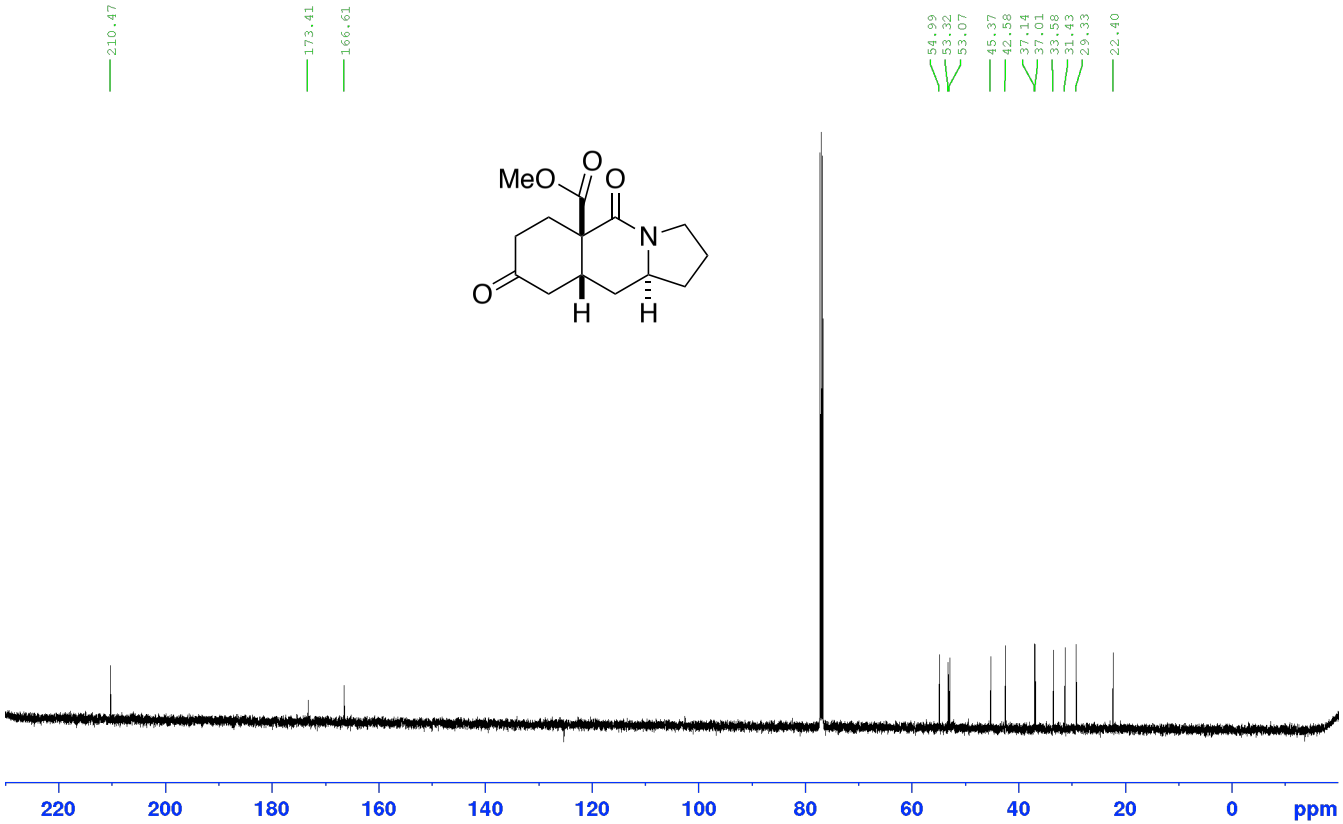
^{1}H NMR of compound **14** (500 MHz, $CDCl_3$)



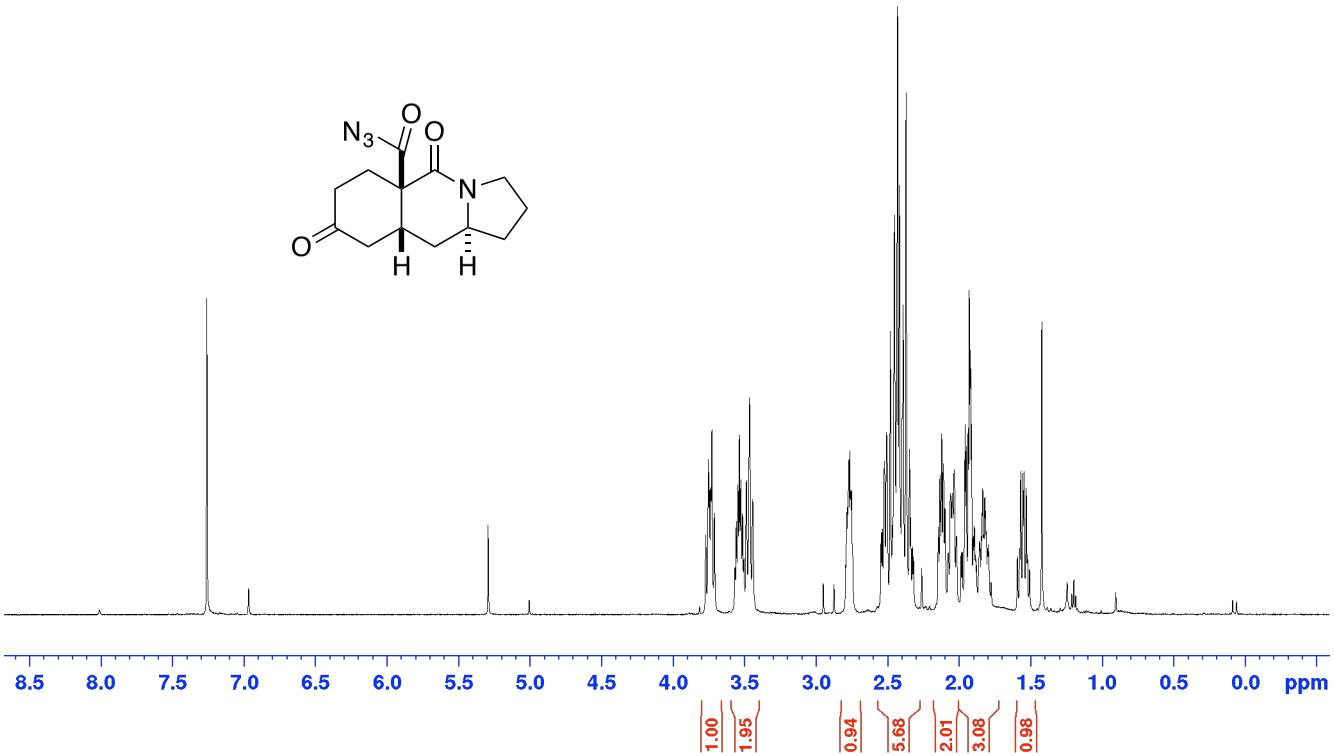
^{13}C NMR of compound **14** (125 MHz, $CDCl_3$)



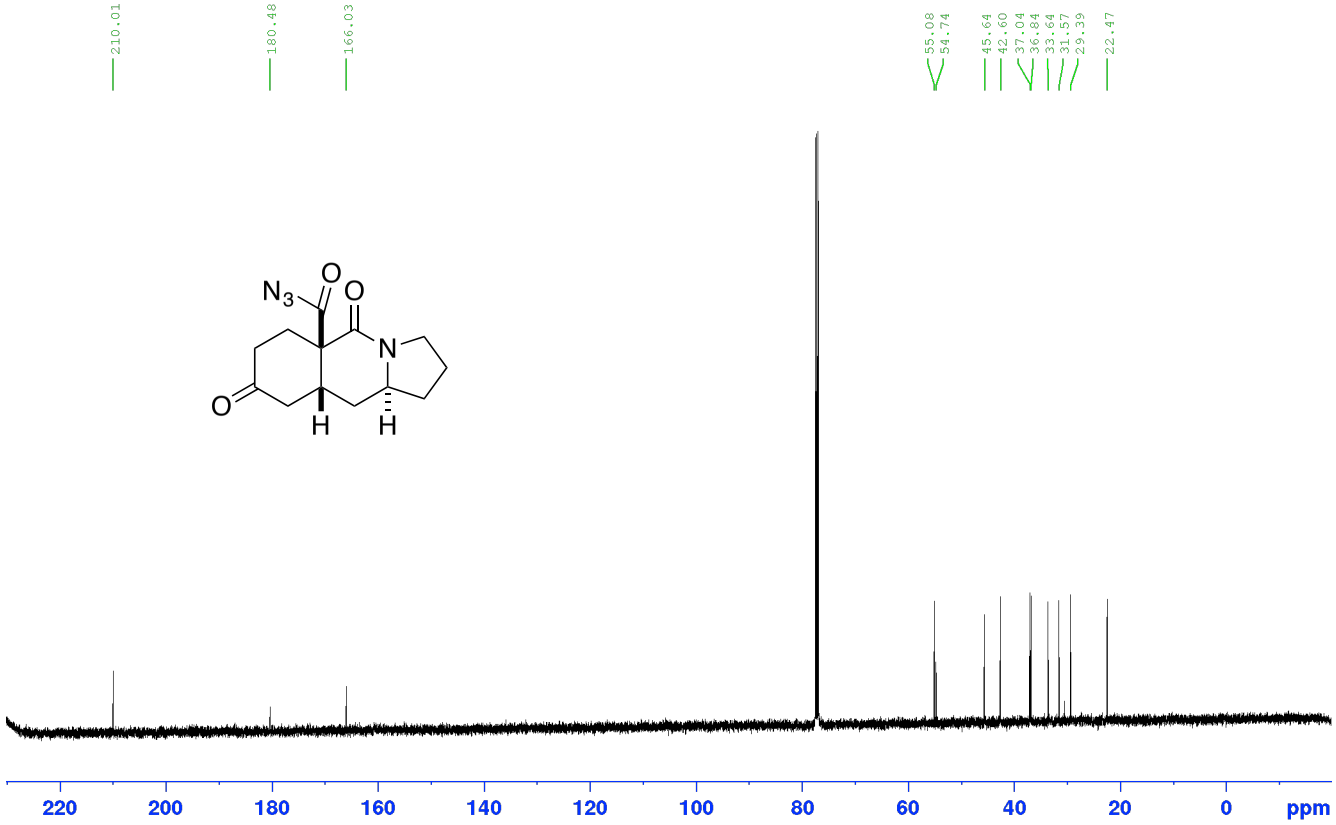
^{1}H NMR of compound **15** (500 MHz, $CDCl_3$)



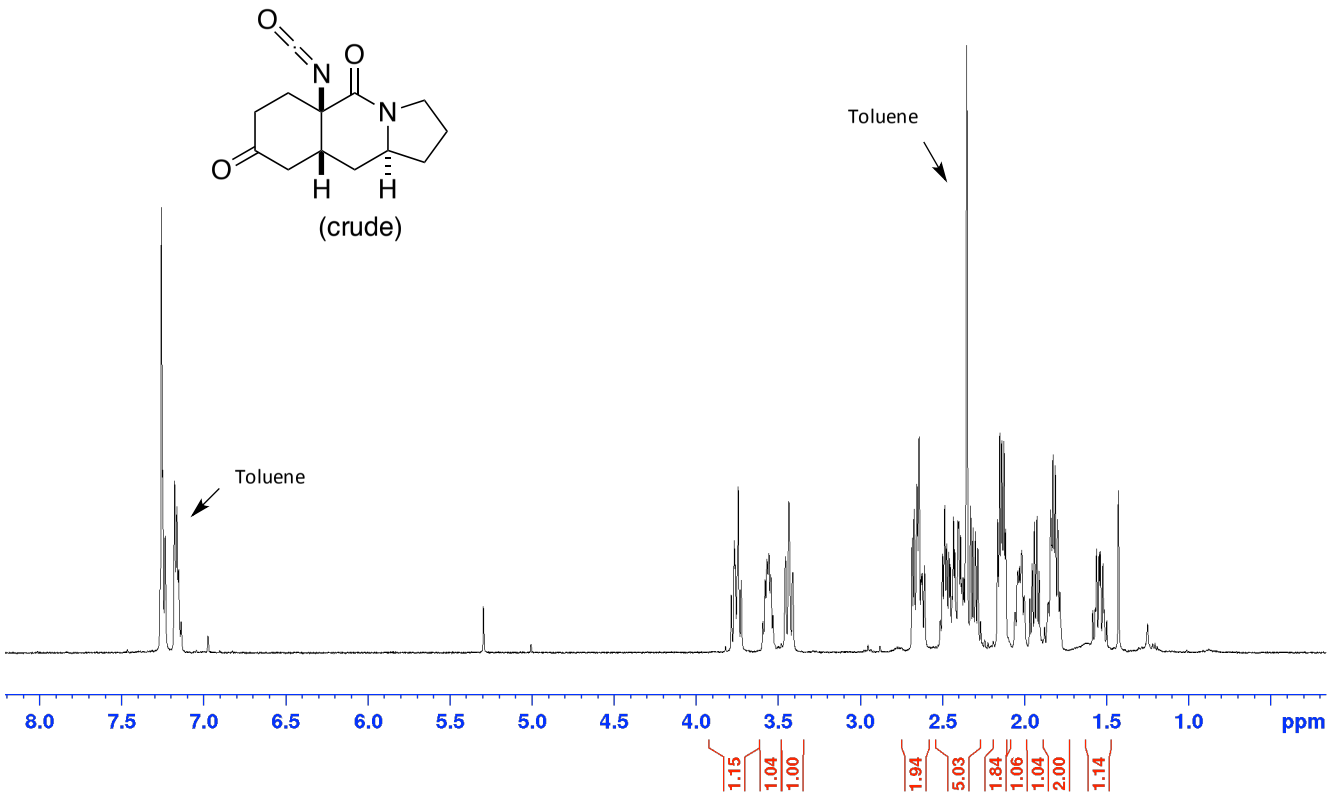
^{13}C NMR of compound **15** (125 MHz, $CDCl_3$)



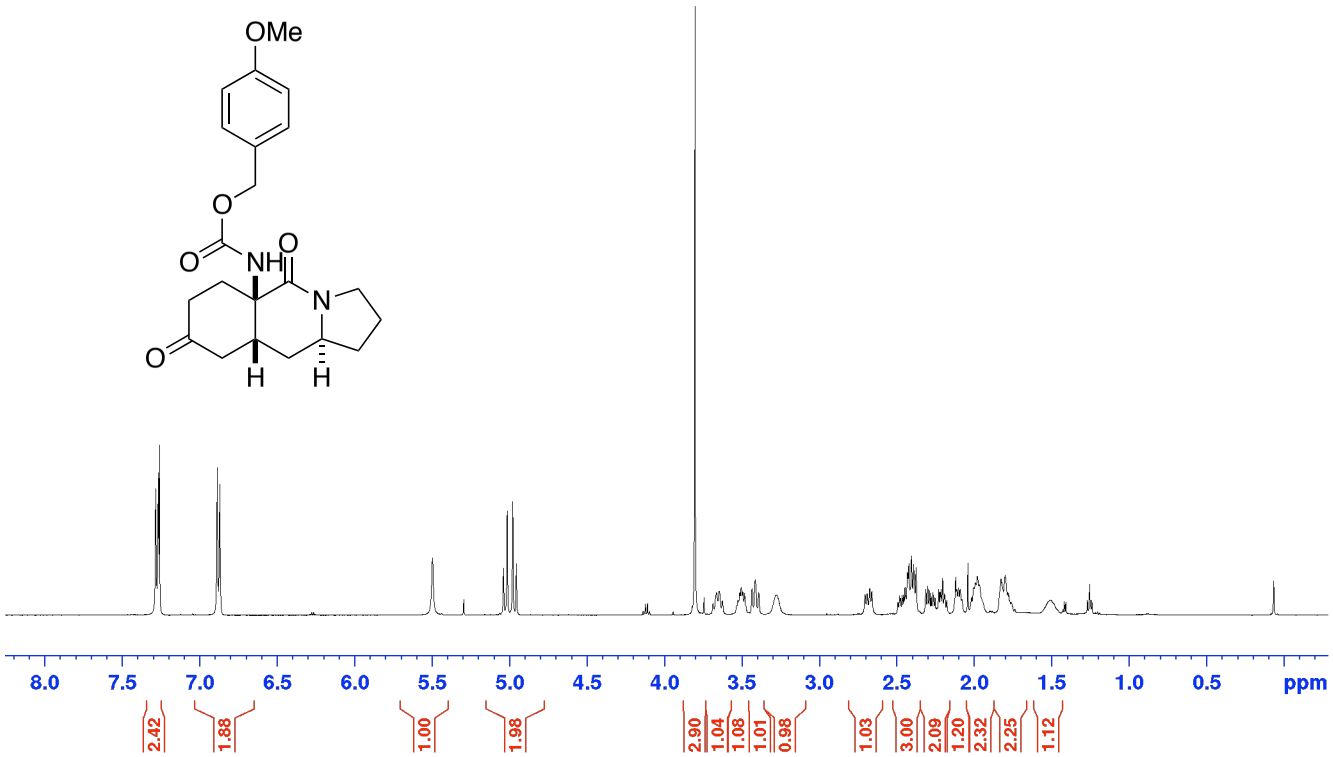
^{1}H NMR of compound **16** (500 MHz, $CDCl_3$)



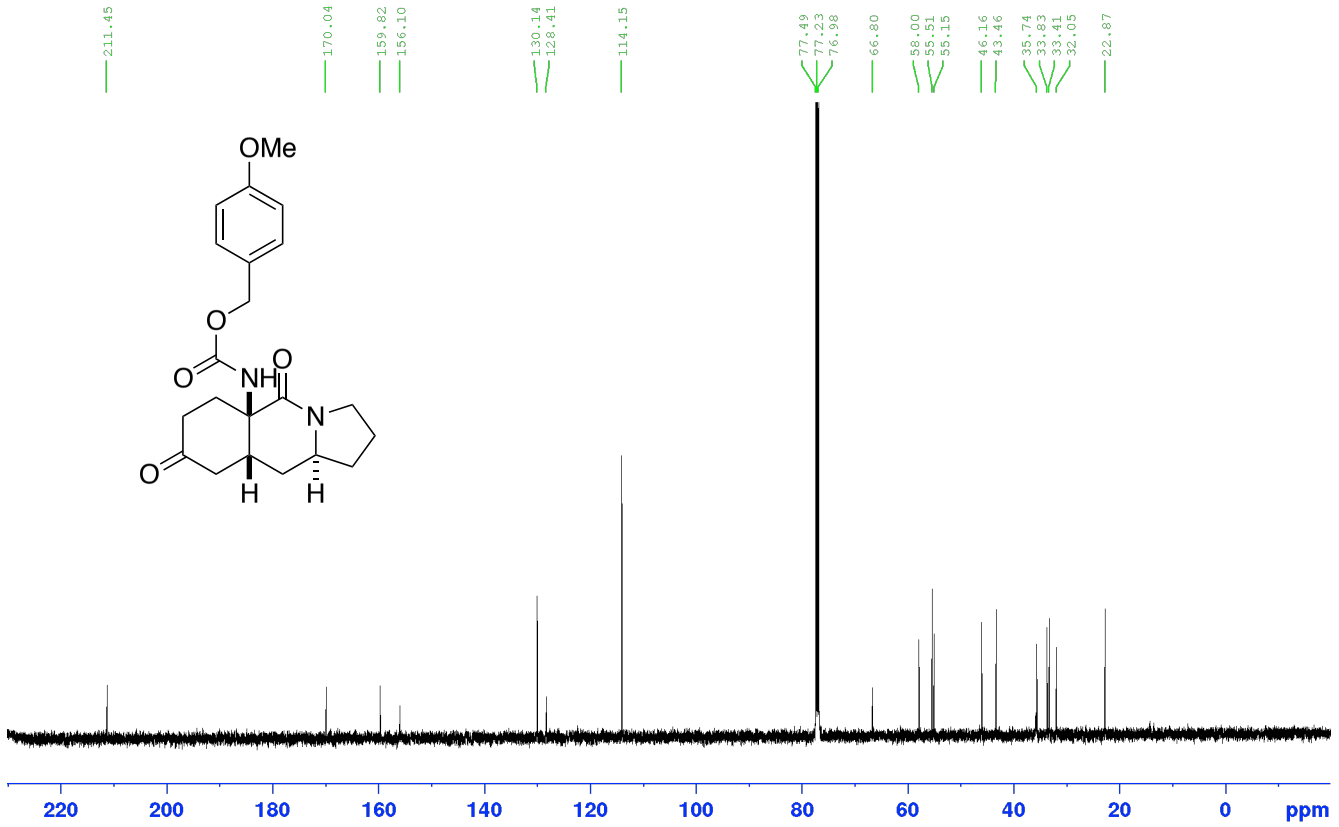
^{13}C NMR of compound **16** (125 MHz, $CDCl_3$)



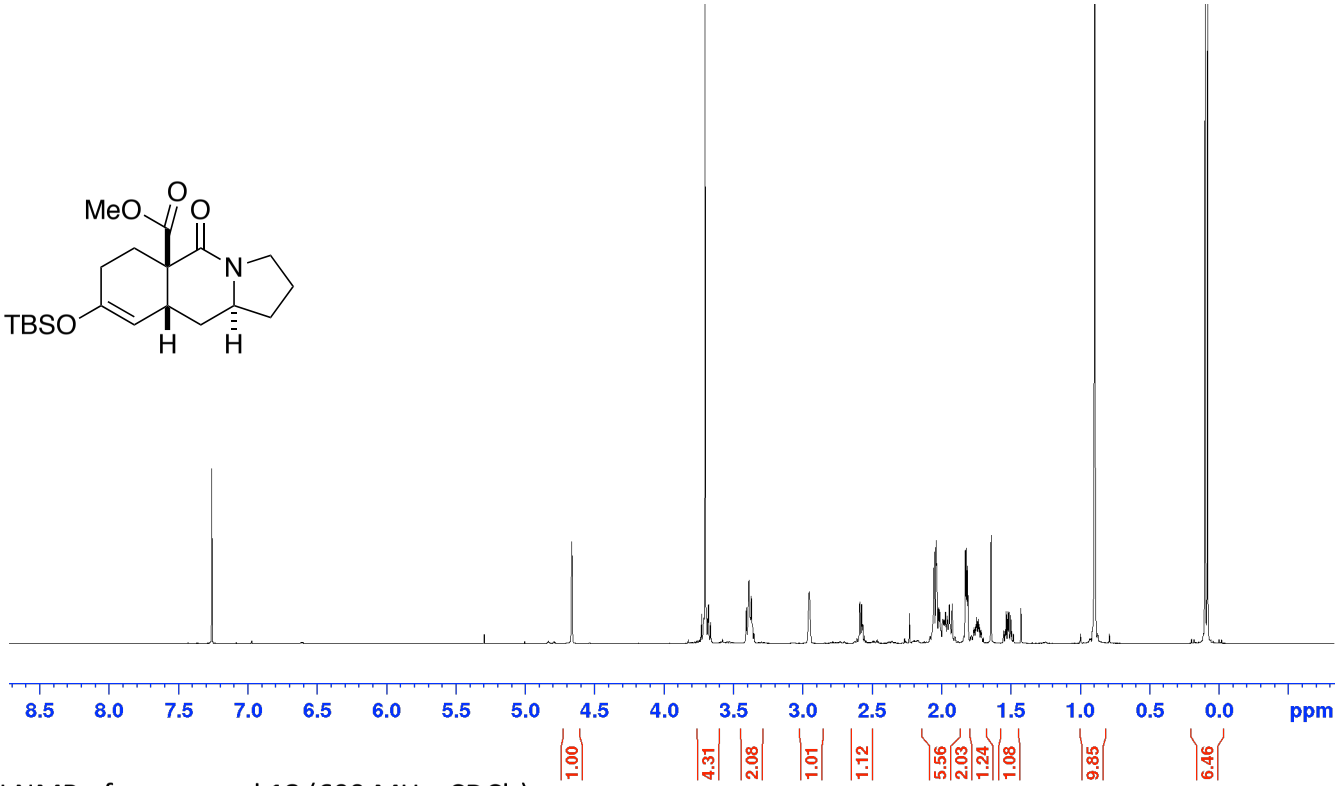
^{1}H NMR of crude isocyanate from thermolysis of **16** (500 MHz, $CDCl_3$)



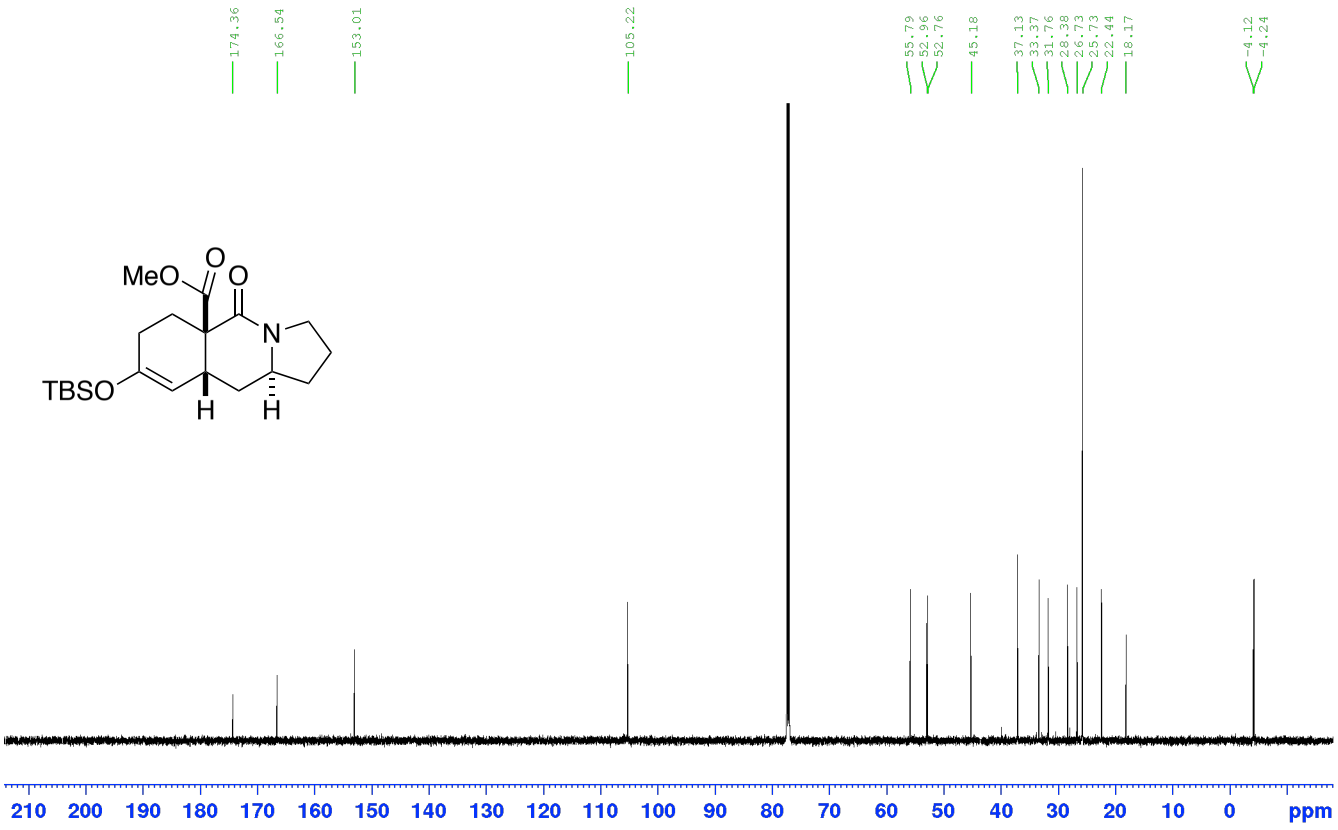
^{1}H NMR of compound **17** (500 MHz, $CDCl_3$)



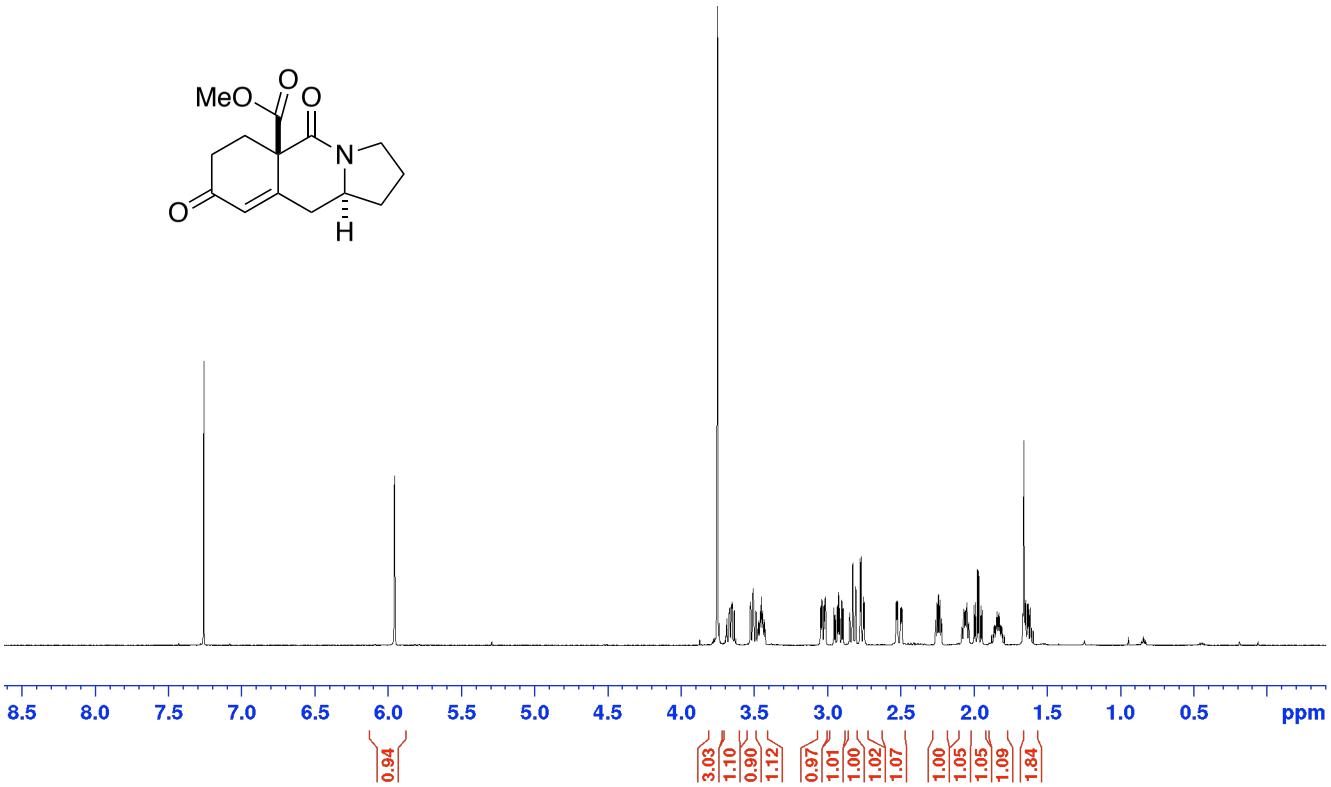
^{13}C NMR of compound **17** (125 MHz, $CDCl_3$)



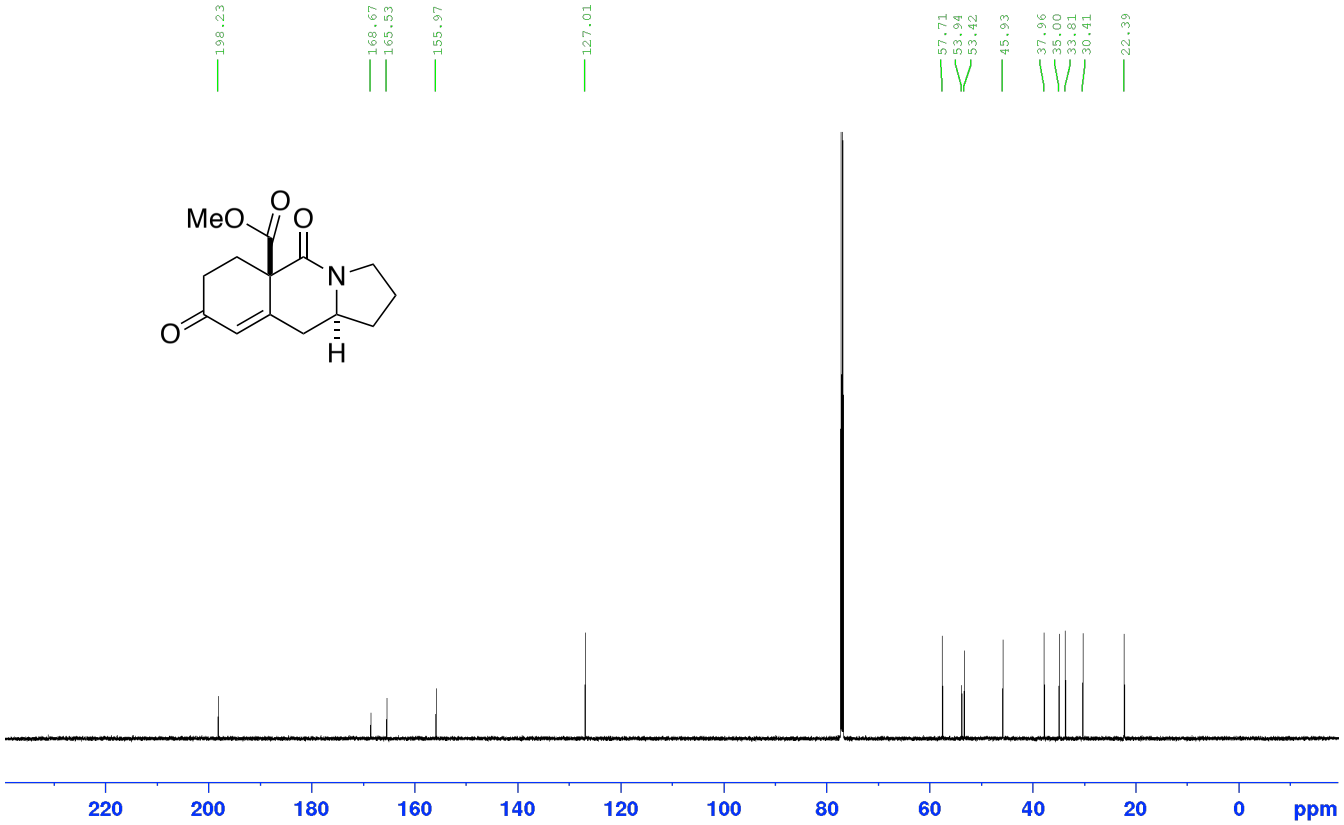
^{1}H NMR of compound **18** (600 MHz, $CDCl_3$)



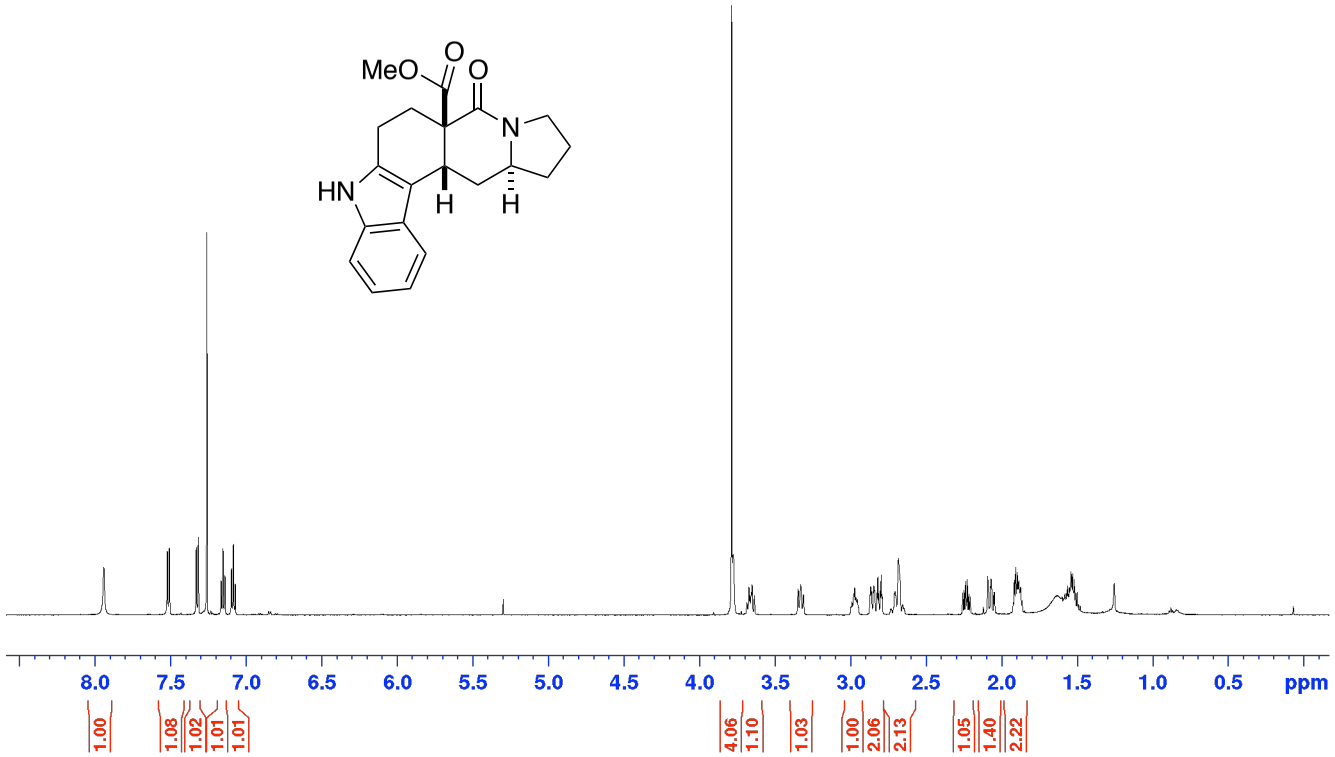
^{13}C NMR of compound **18** (150 MHz, $CDCl_3$)



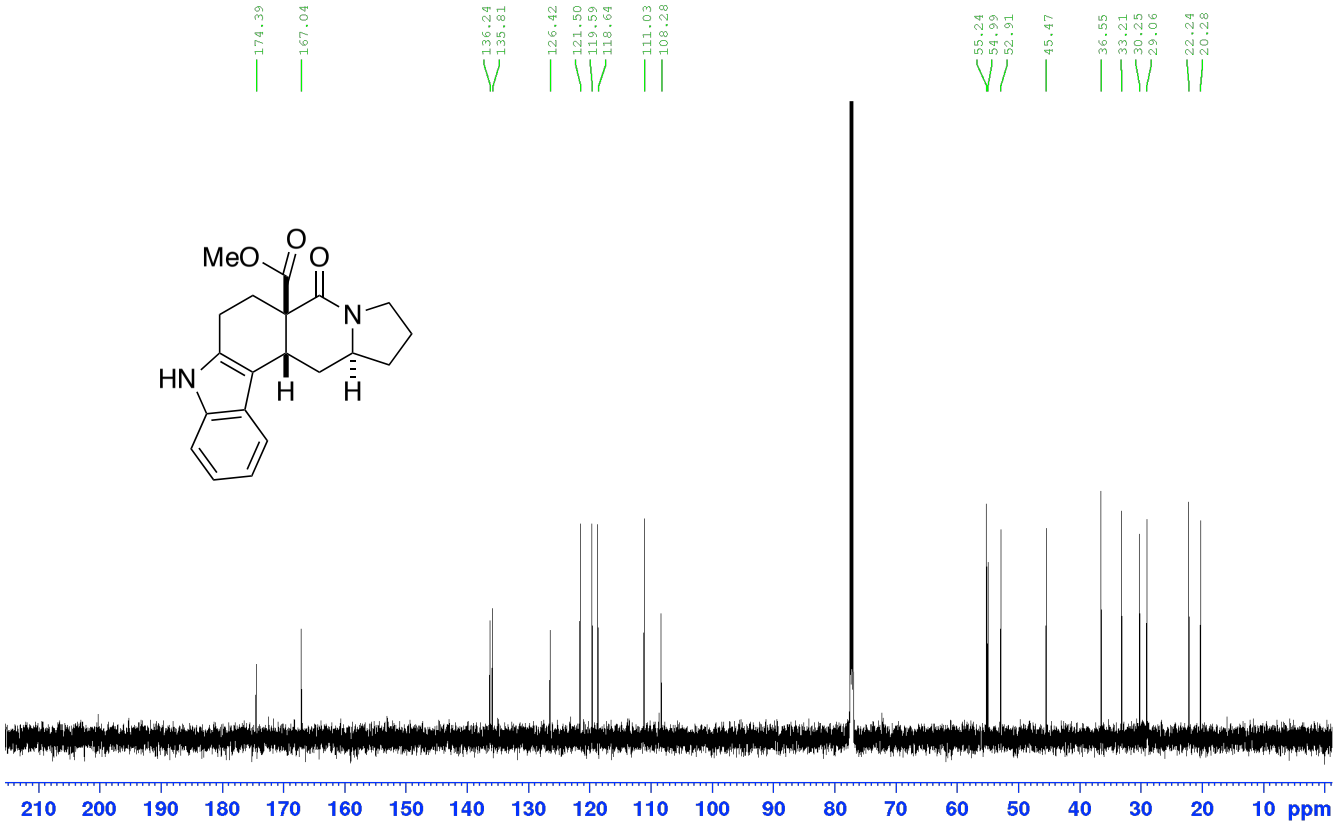
^{1}H NMR of compound **19** (600 MHz, $CDCl_3$)



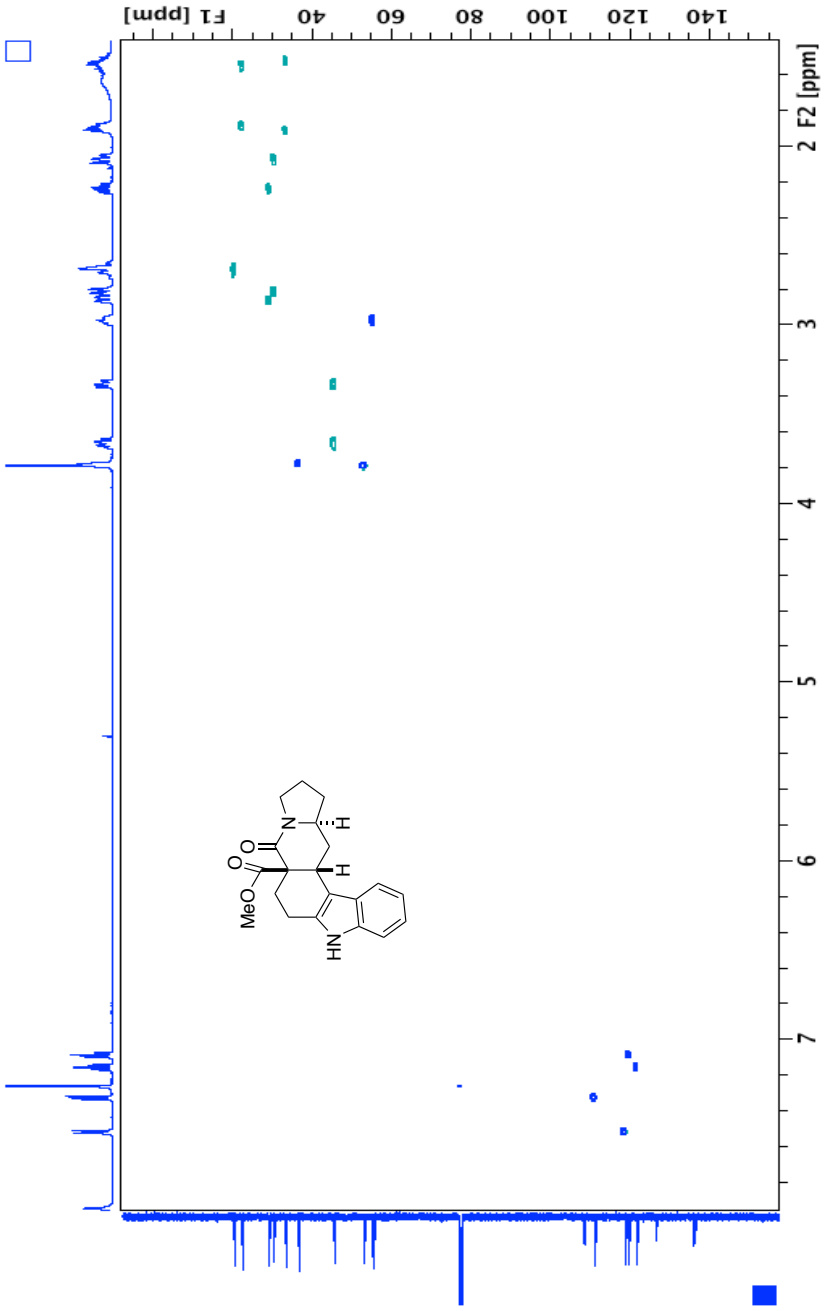
^{13}C NMR of compound **19** (150 MHz, $CDCl_3$)



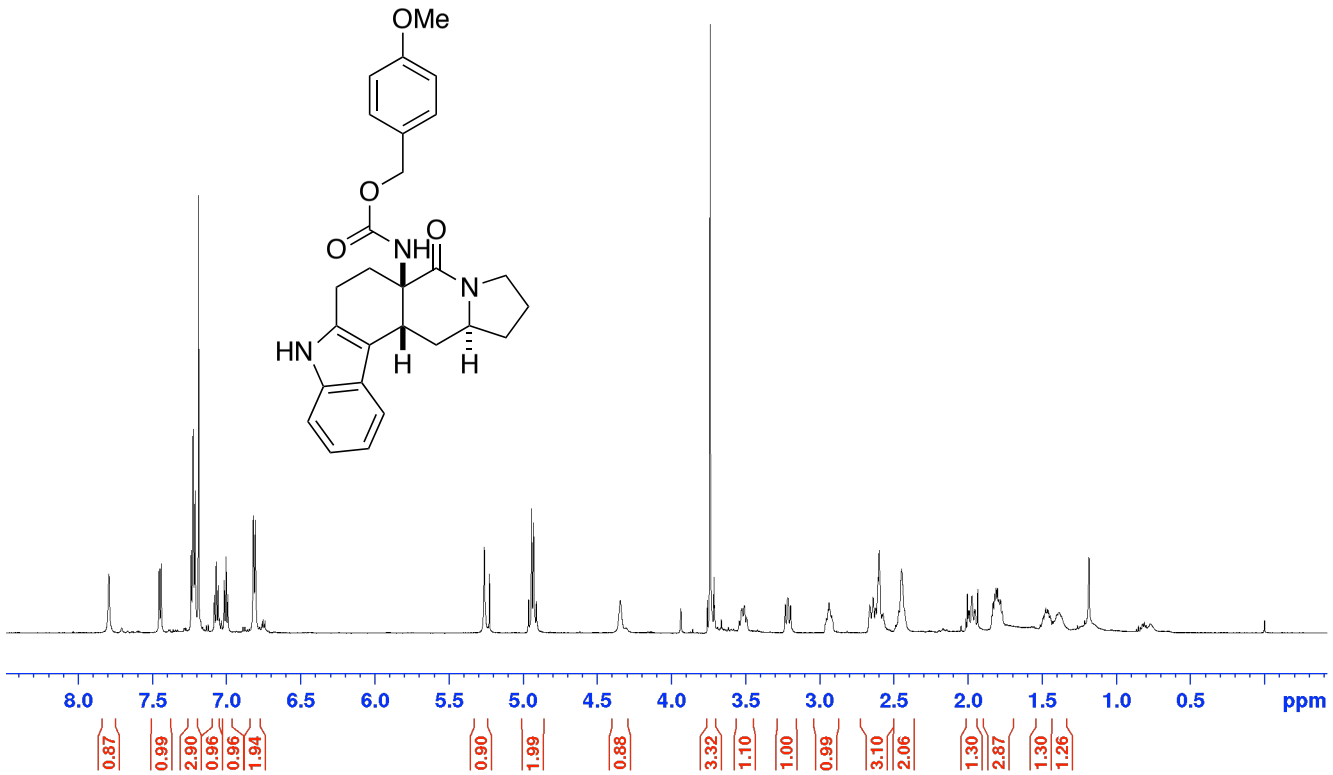
^{1}H NMR of compound **23** (600 MHz, $CDCl_3$)



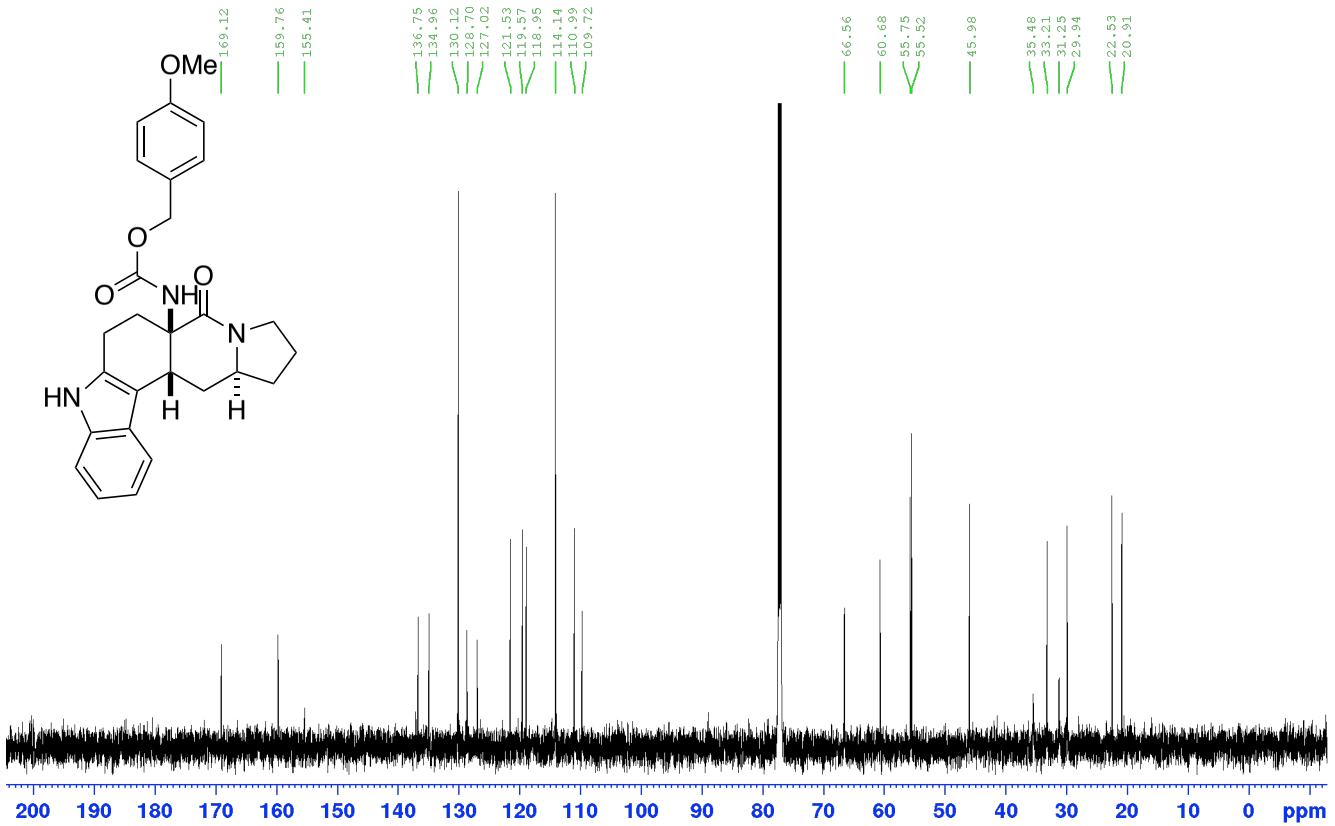
^{13}C NMR of compound **23** (150 MHz, $CDCl_3$)



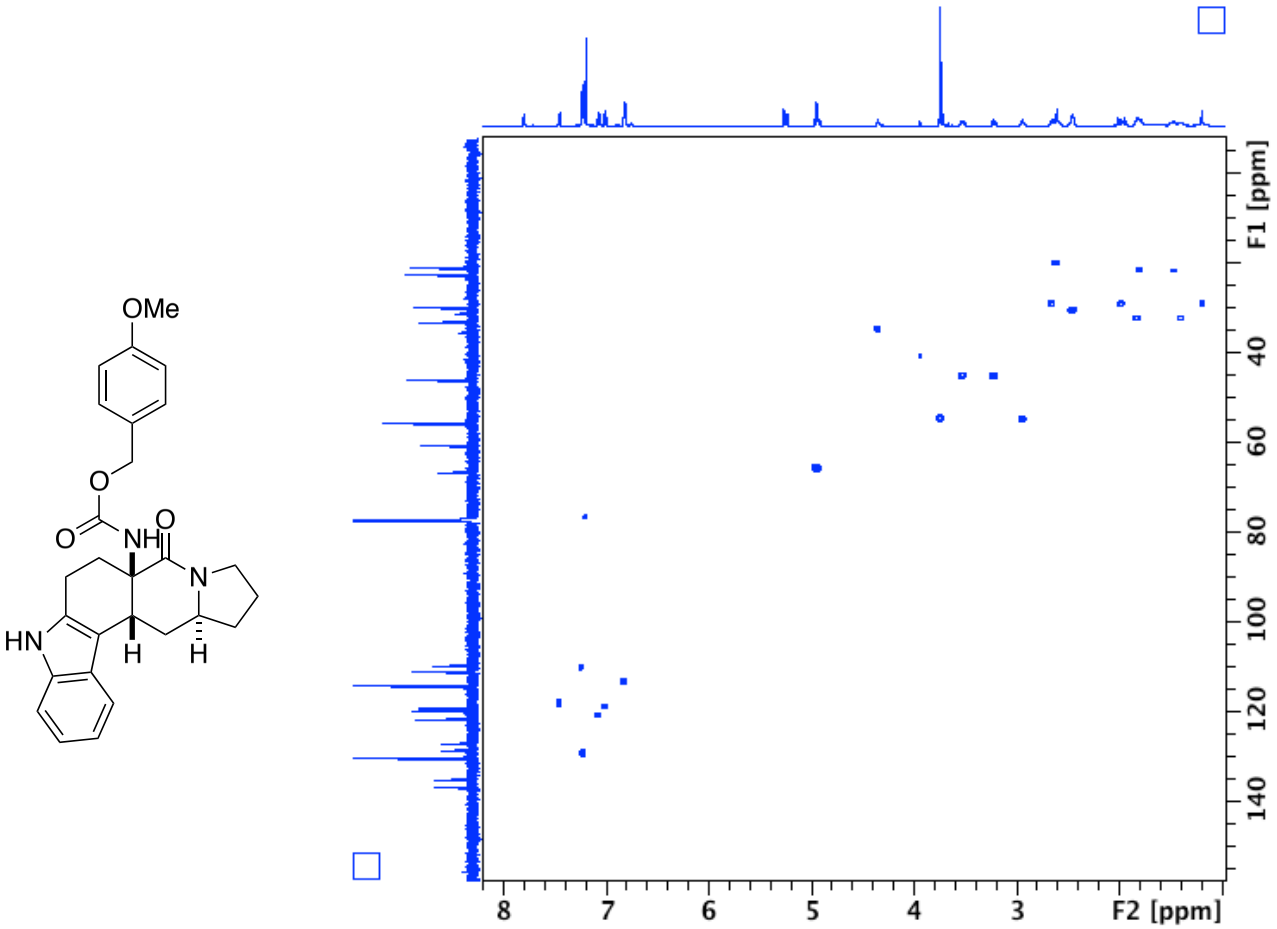
HSQC spectrum of compound **23** (150 MHz, $CDCl_3$)



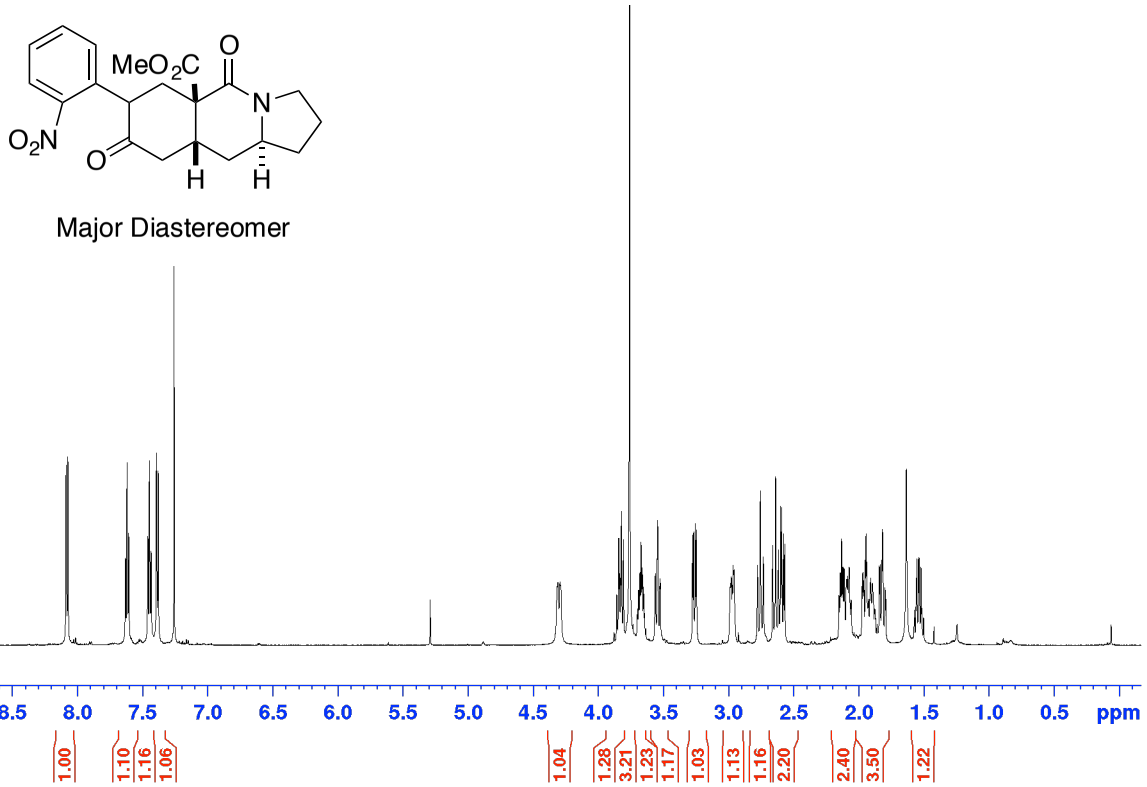
^{1}H NMR of compound **25** (600 MHz, $CDCl_3$)



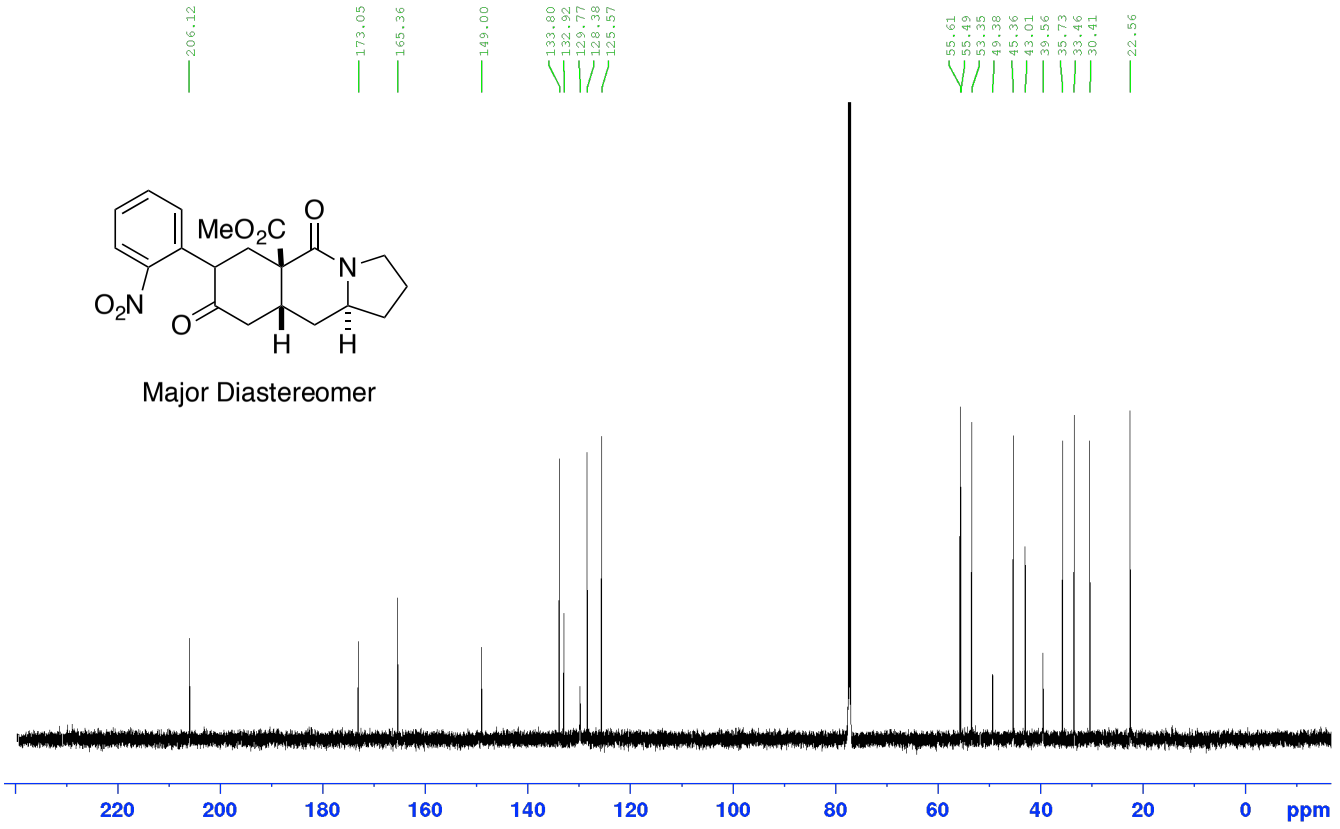
^{13}C NMR of compound **25** (150 MHz, $CDCl_3$)



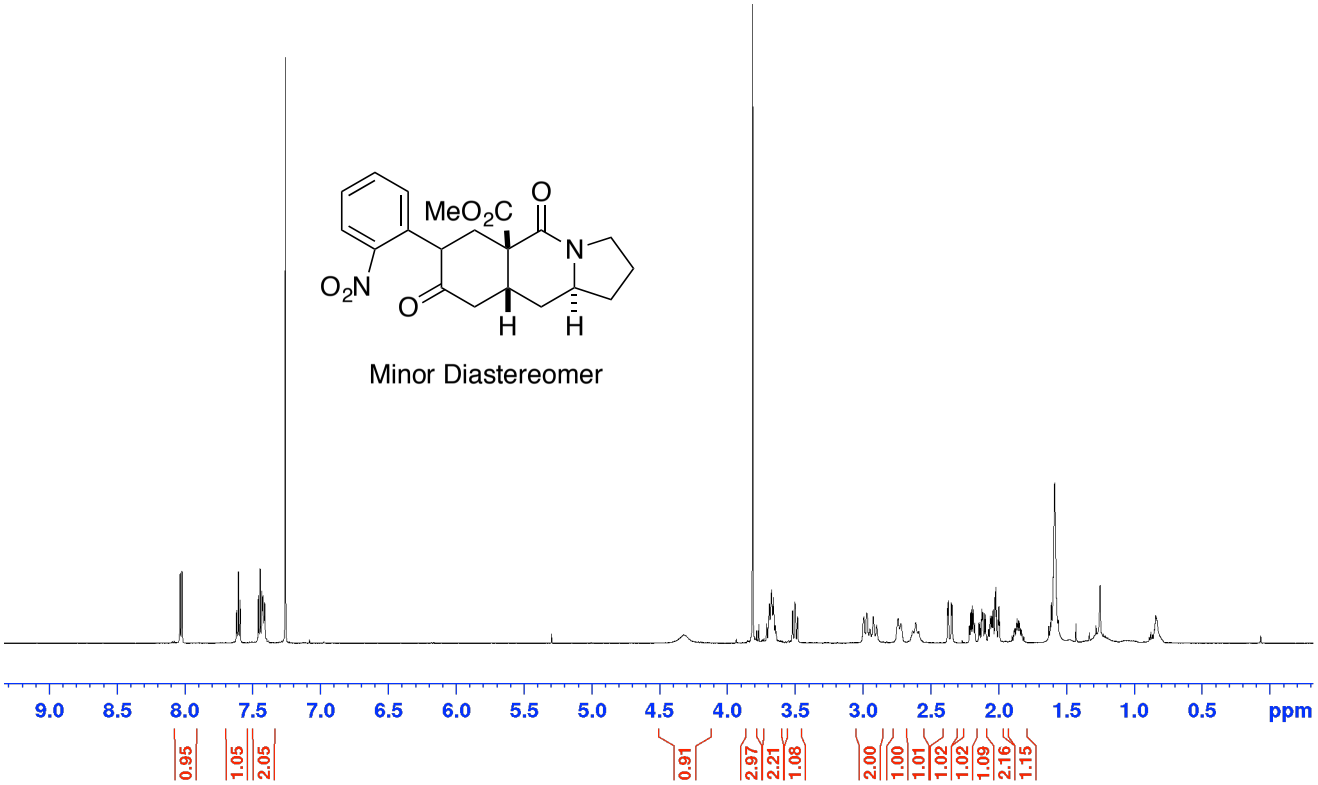
HSQC spectrum of compound **25** (150 MHz, $CDCl_3$)



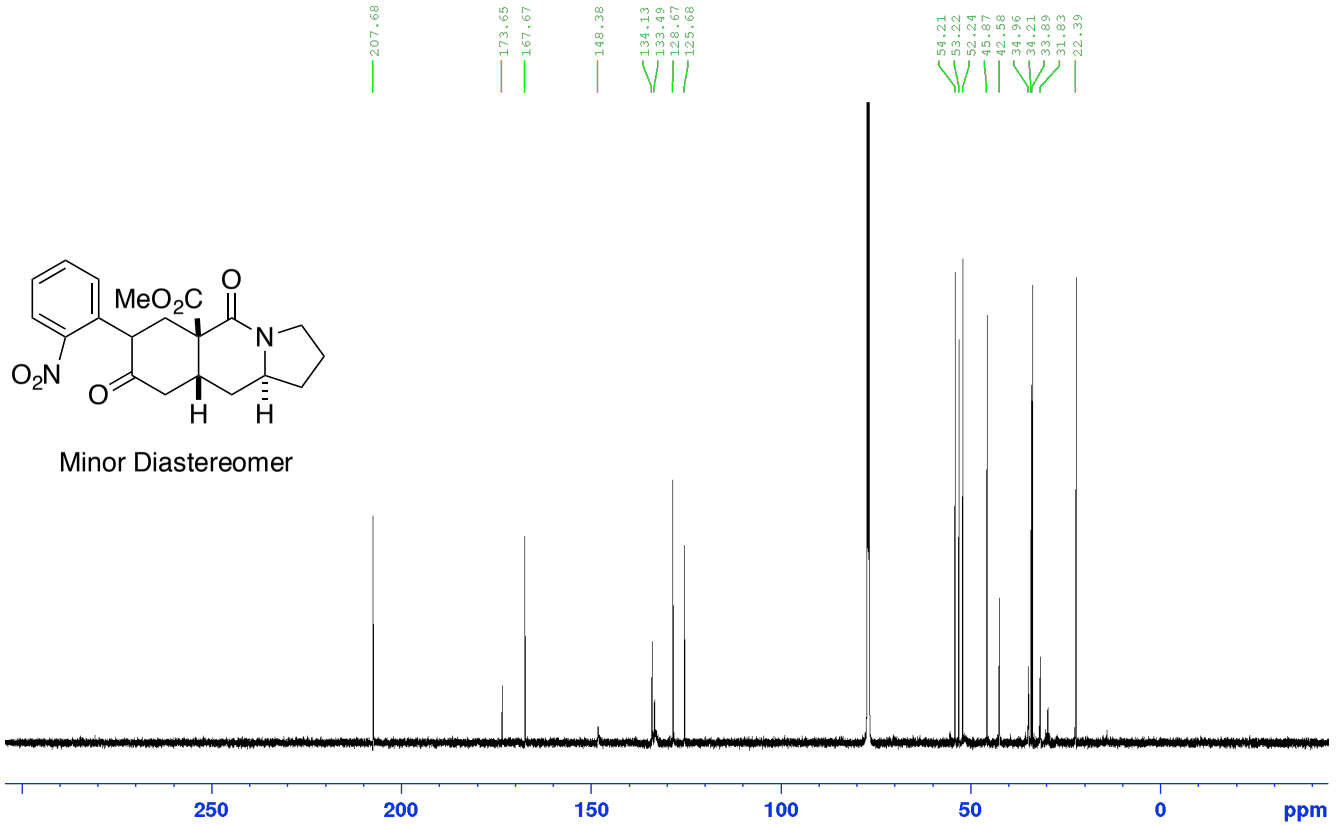
^{1}H NMR of compound **27** – Major Diastereomer (600 MHz, $CDCl_3$)



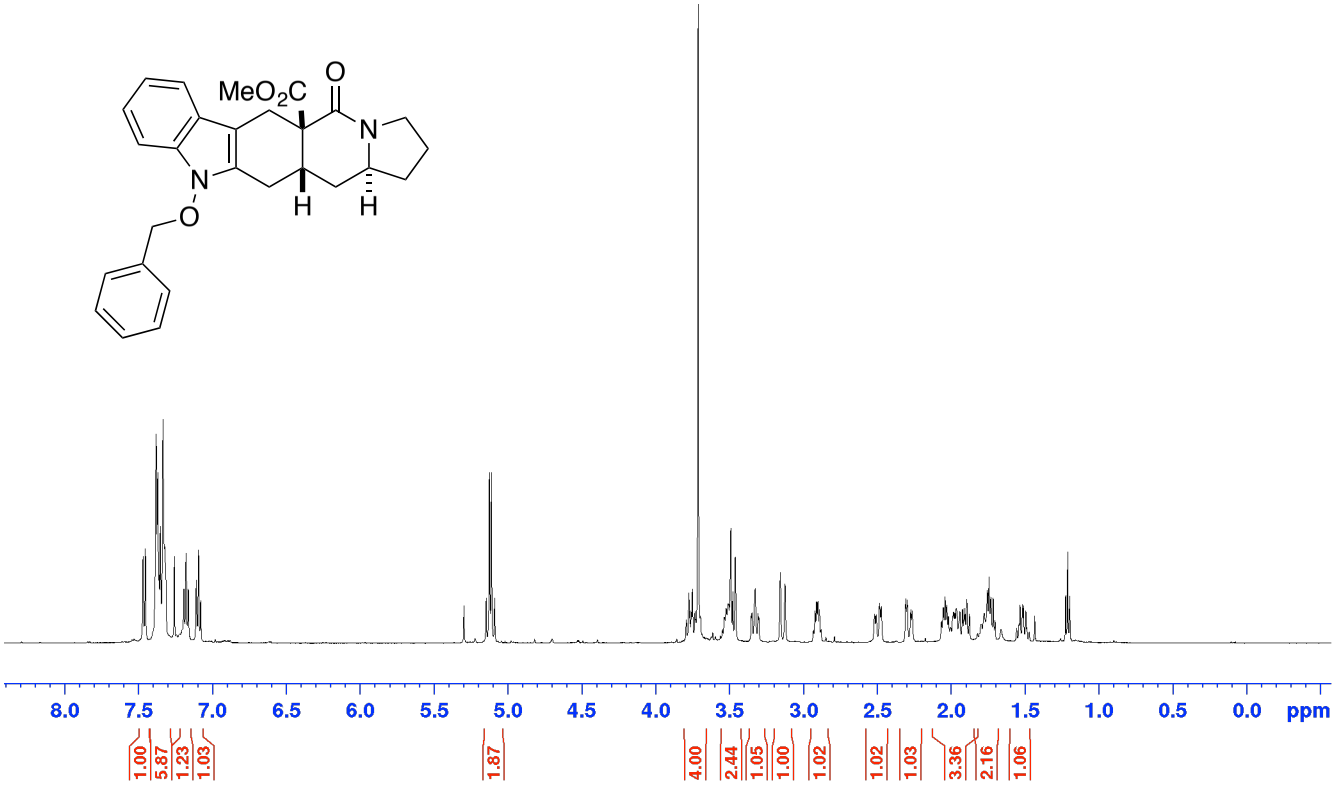
^{13}C NMR of compound **27** – Major Diastereomer (150 MHz, $CDCl_3$)



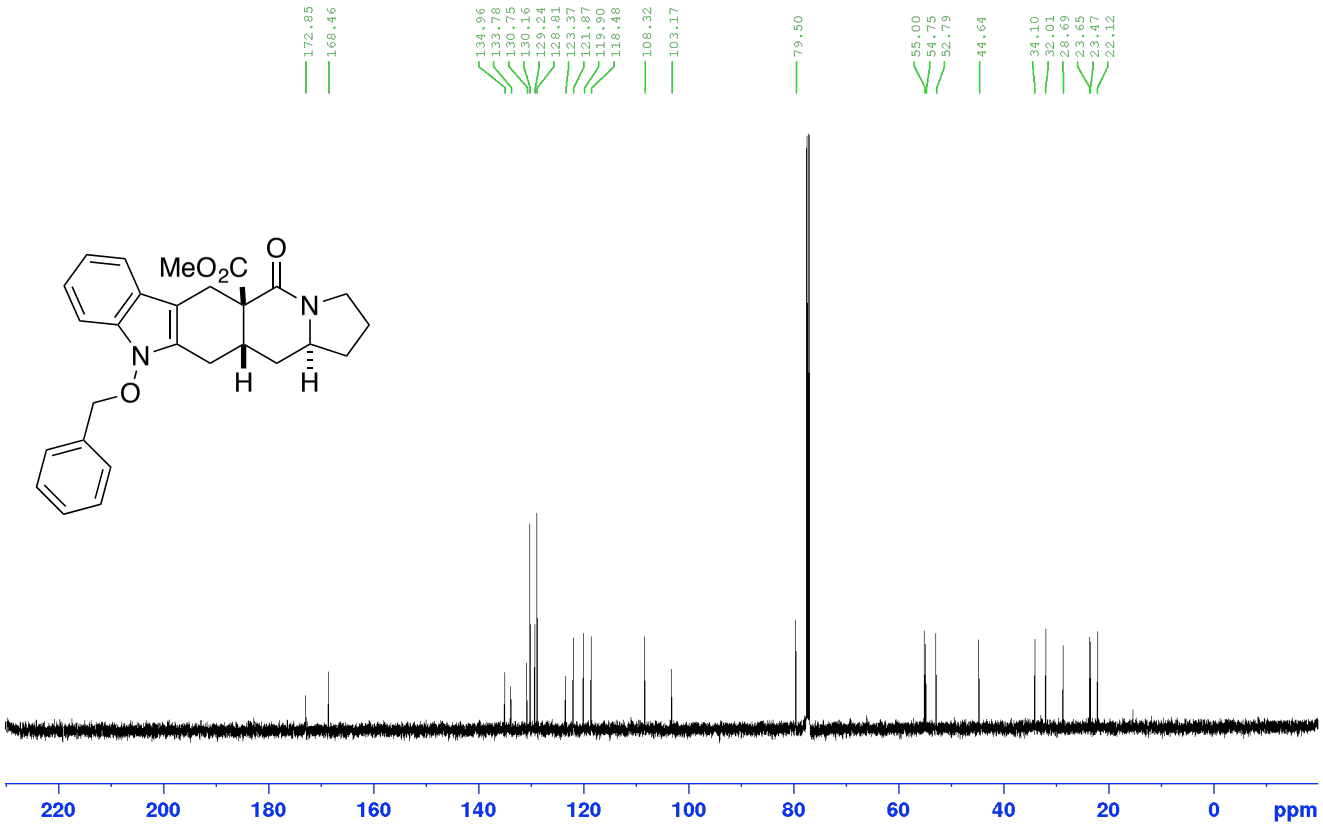
^{1}H NMR of compound **27** – Minor Diastereomer (600 MHz, $CDCl_3$)



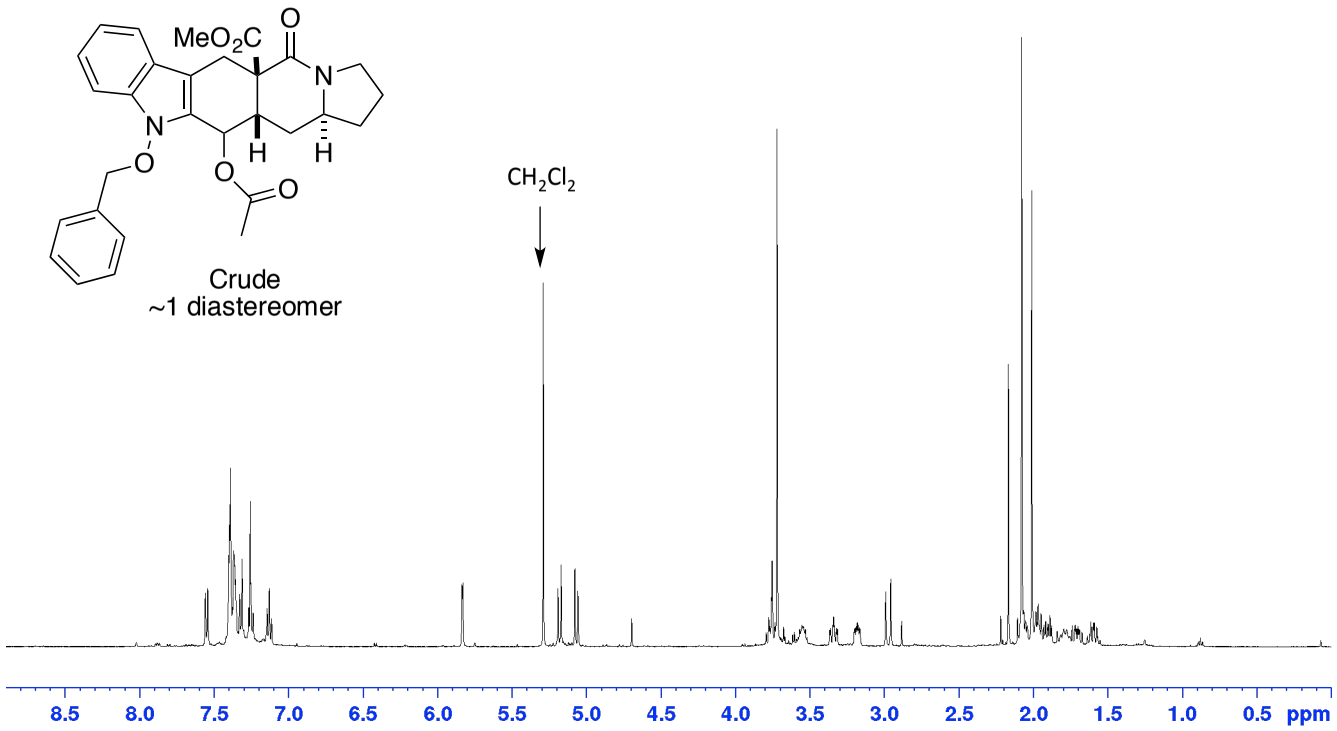
^{13}C NMR of compound **27** – Minor Diastereomer (150 MHz, $CDCl_3$)



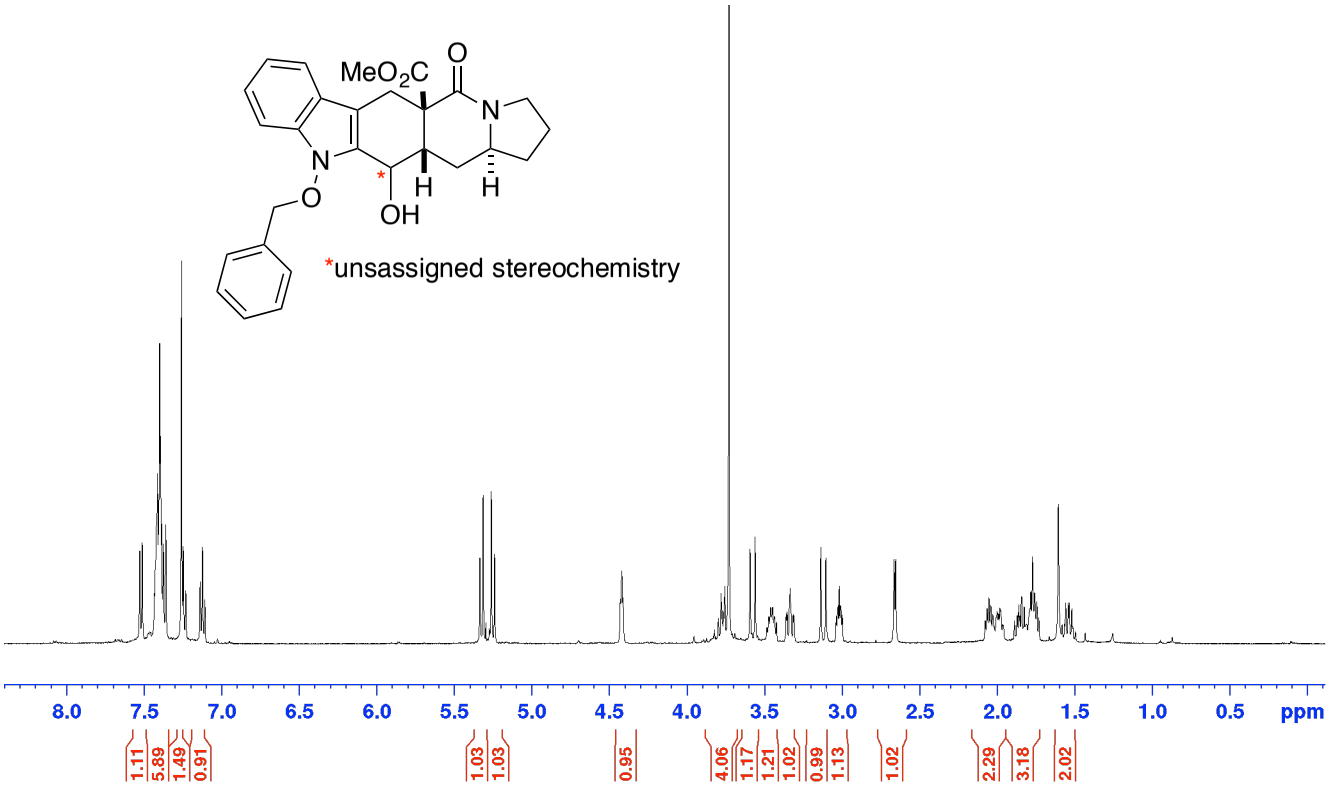
^{1}H NMR of compound **28** (500 MHz, $CDCl_3$)



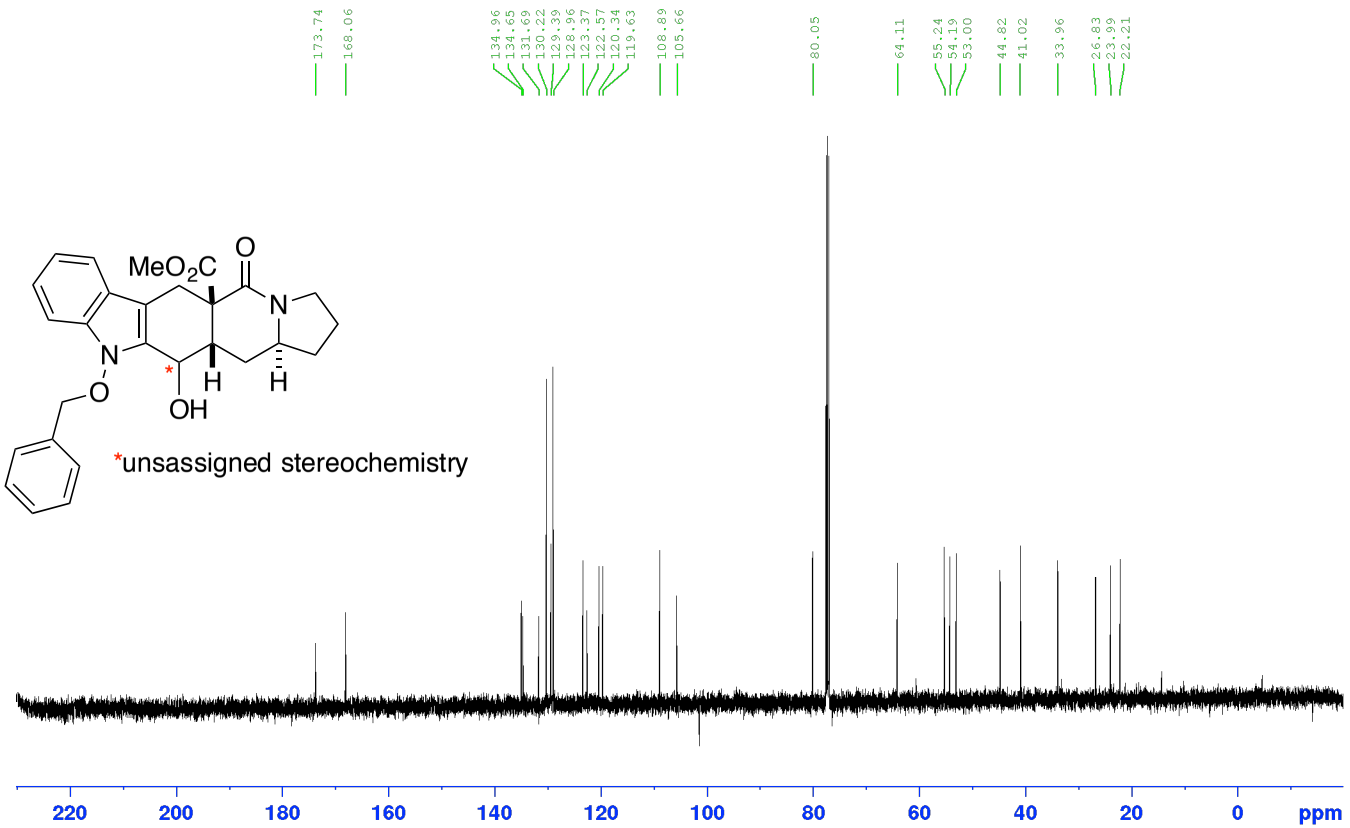
^{13}C NMR of compound **28** (125 MHz, $CDCl_3$)



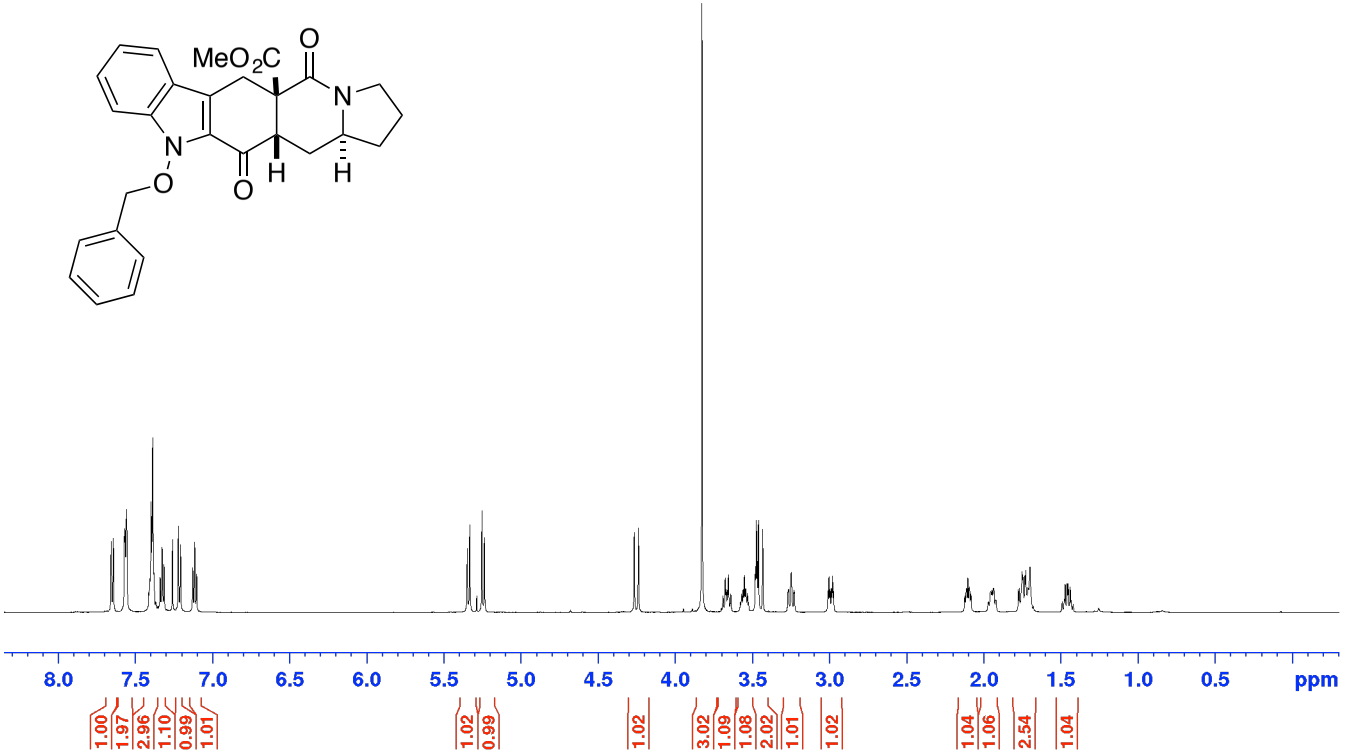
^{1}H NMR of compound crude compound **29** (500 MHz, $CDCl_3$)



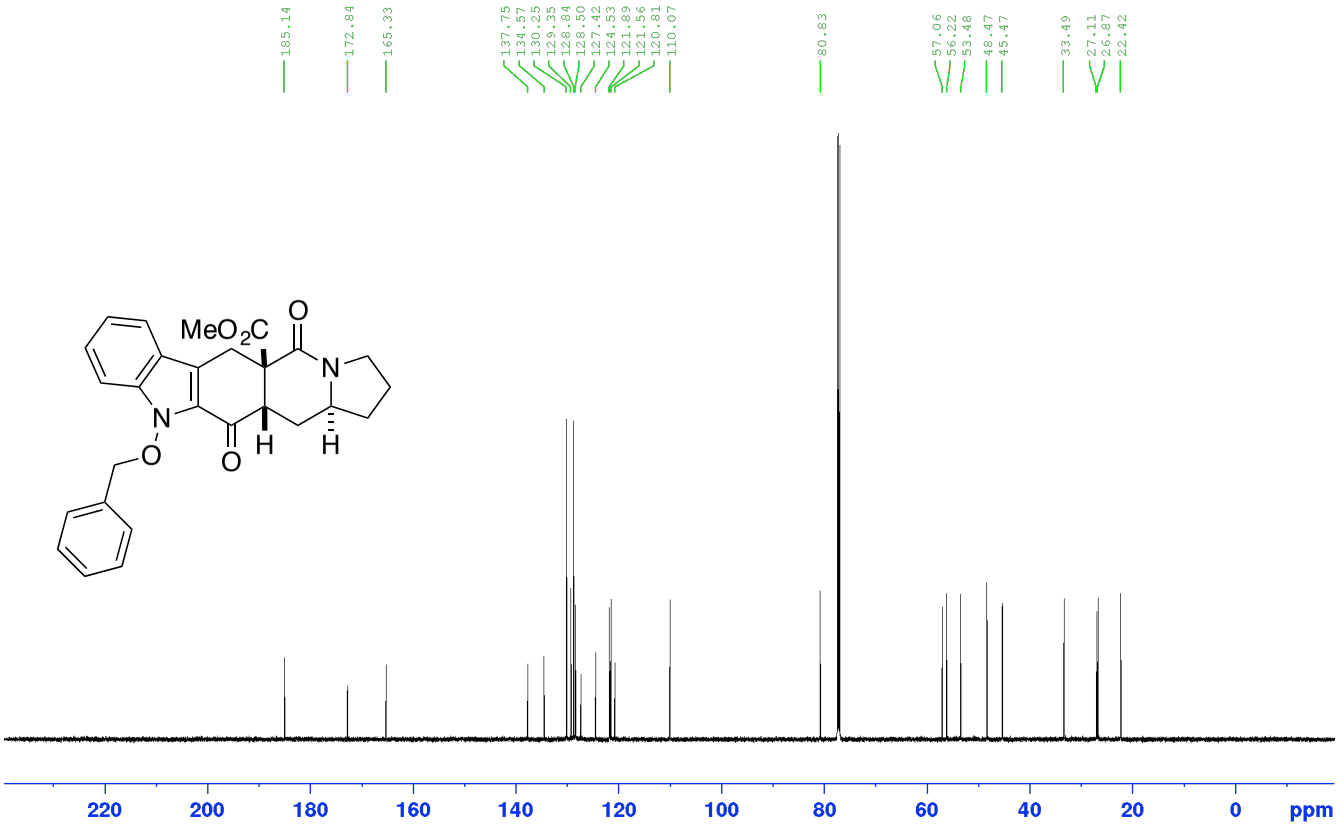
^{1}H NMR of compound **30** (500 MHz, $CDCl_3$)



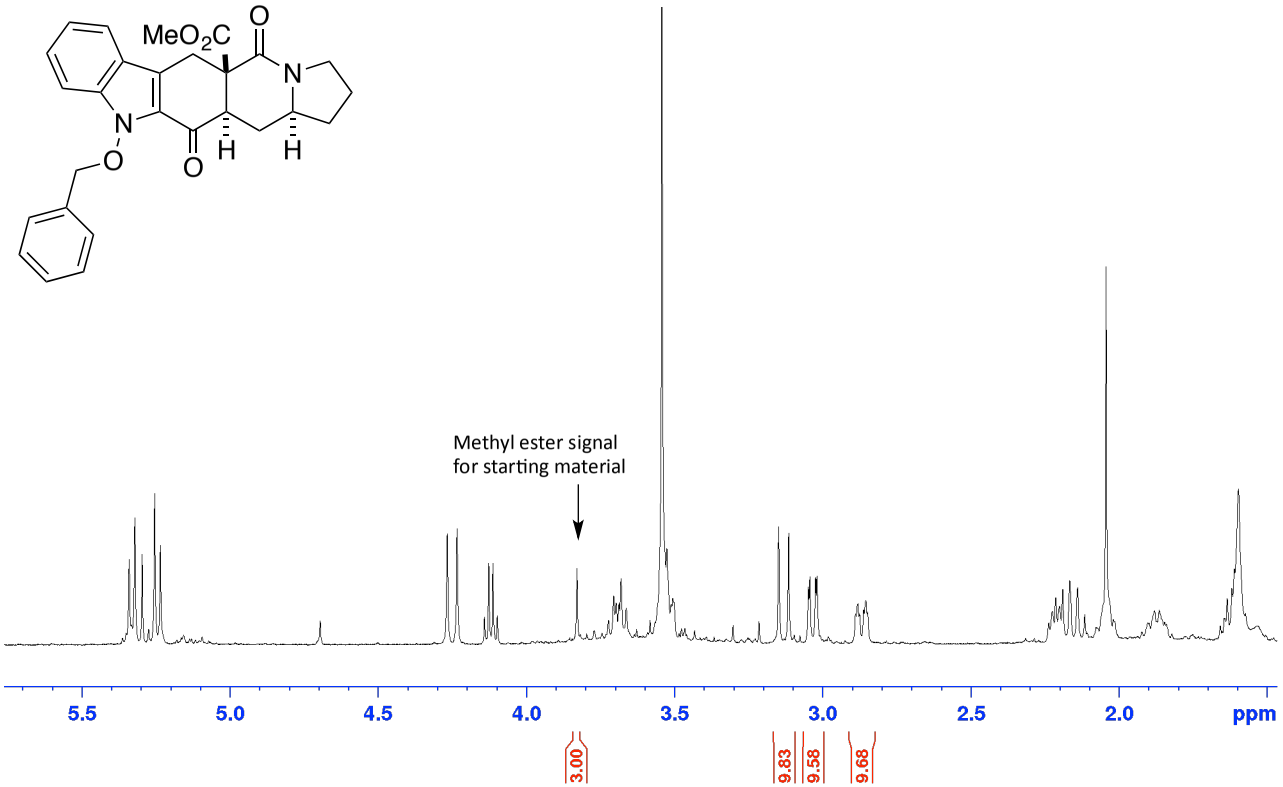
^{13}C NMR of compound **30** (125 MHz, $CDCl_3$)



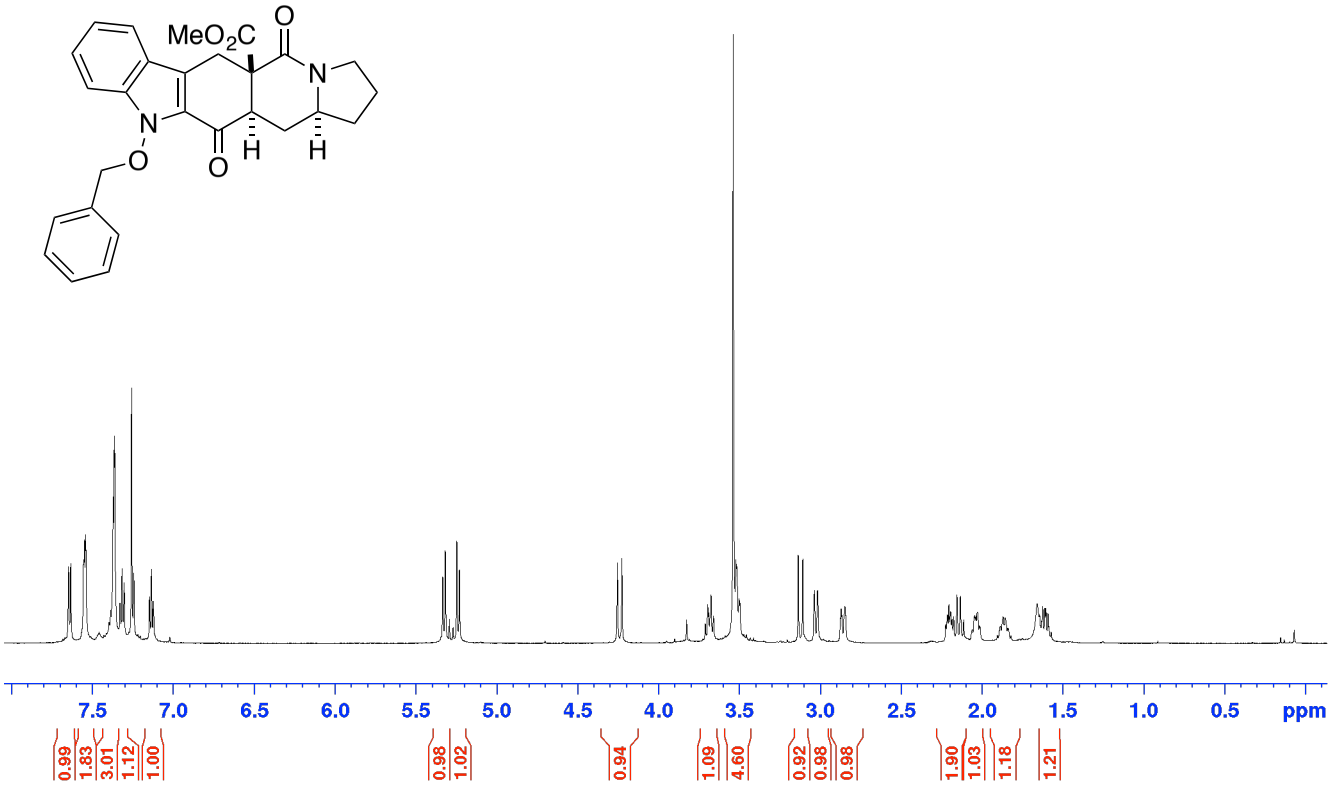
^{1}H NMR of compound **31** (600 MHz, $CDCl_3$)



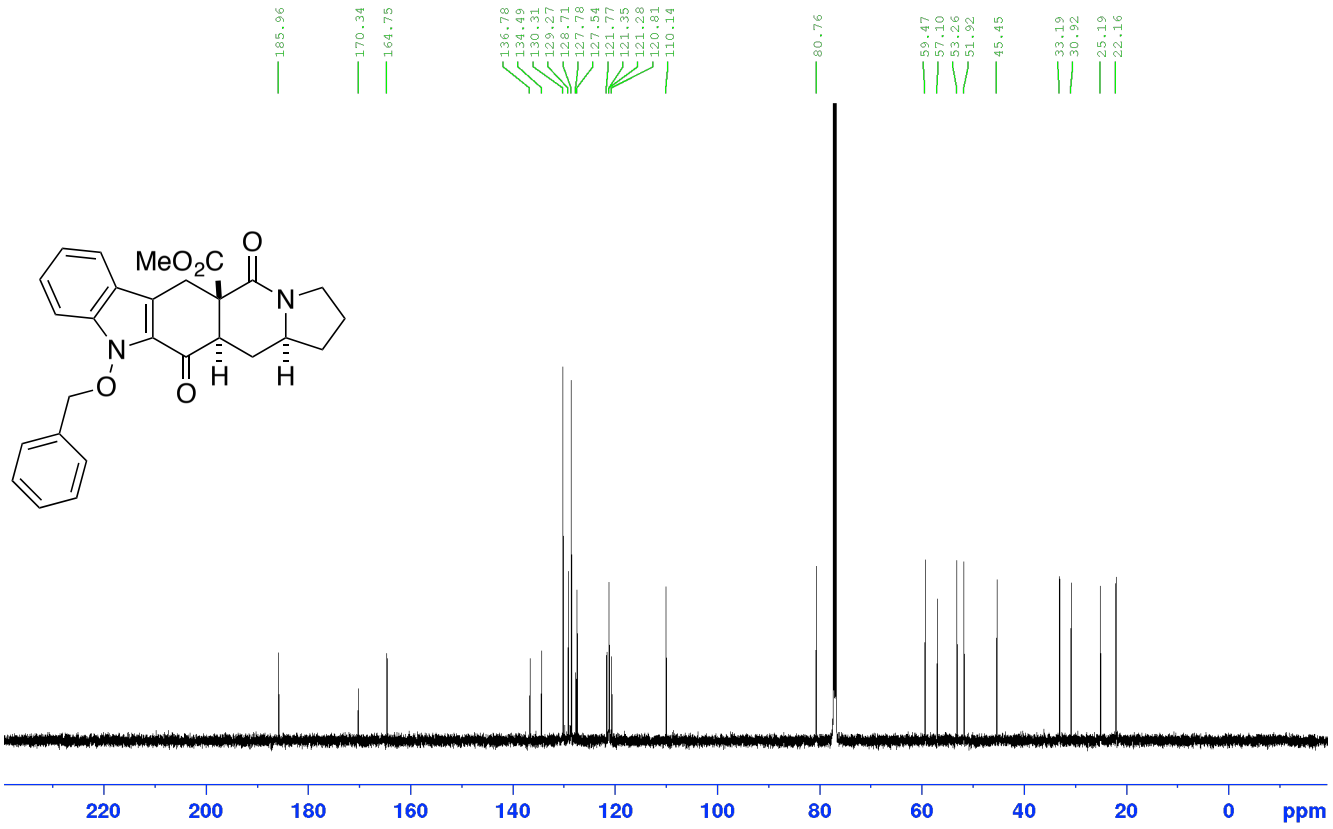
^{13}C NMR of compound **31** (150 MHz, $CDCl_3$)



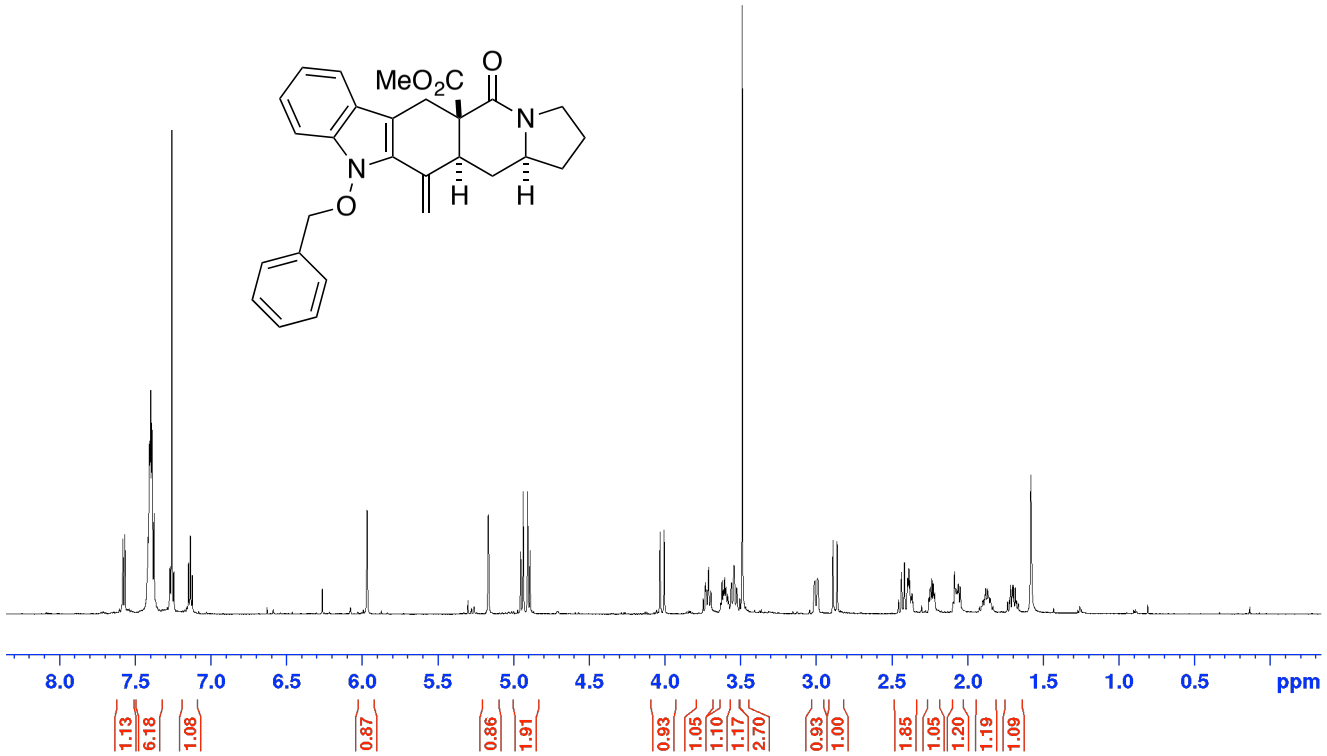
^{1}H NMR of crude compound **32** (500 MHz, $CDCl_3$) showing ~9.5:1 *dr*



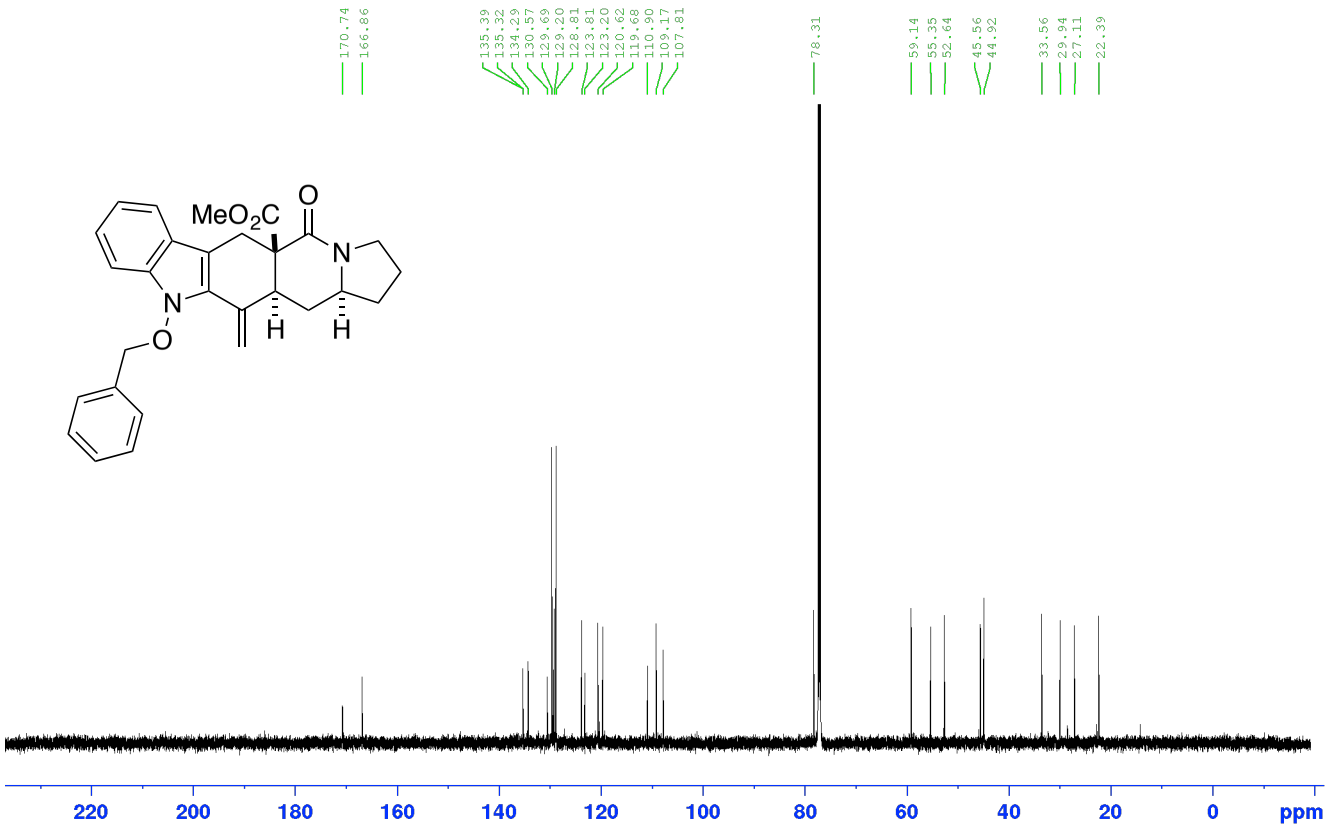
^{1}H NMR of compound **32** (600 MHz, $CDCl_3$)



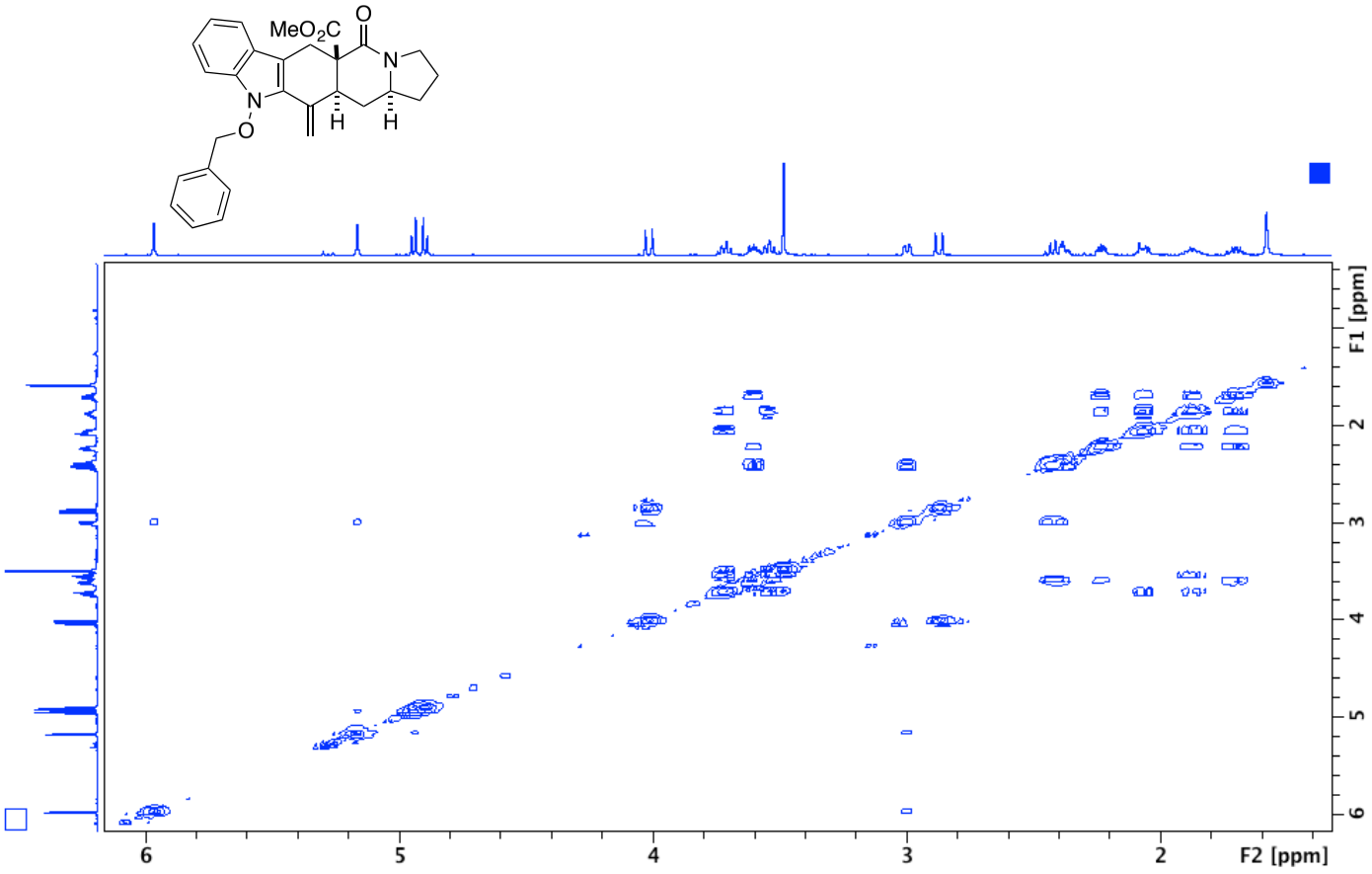
^{13}C NMR of compound **32** (150 MHz, $CDCl_3$)



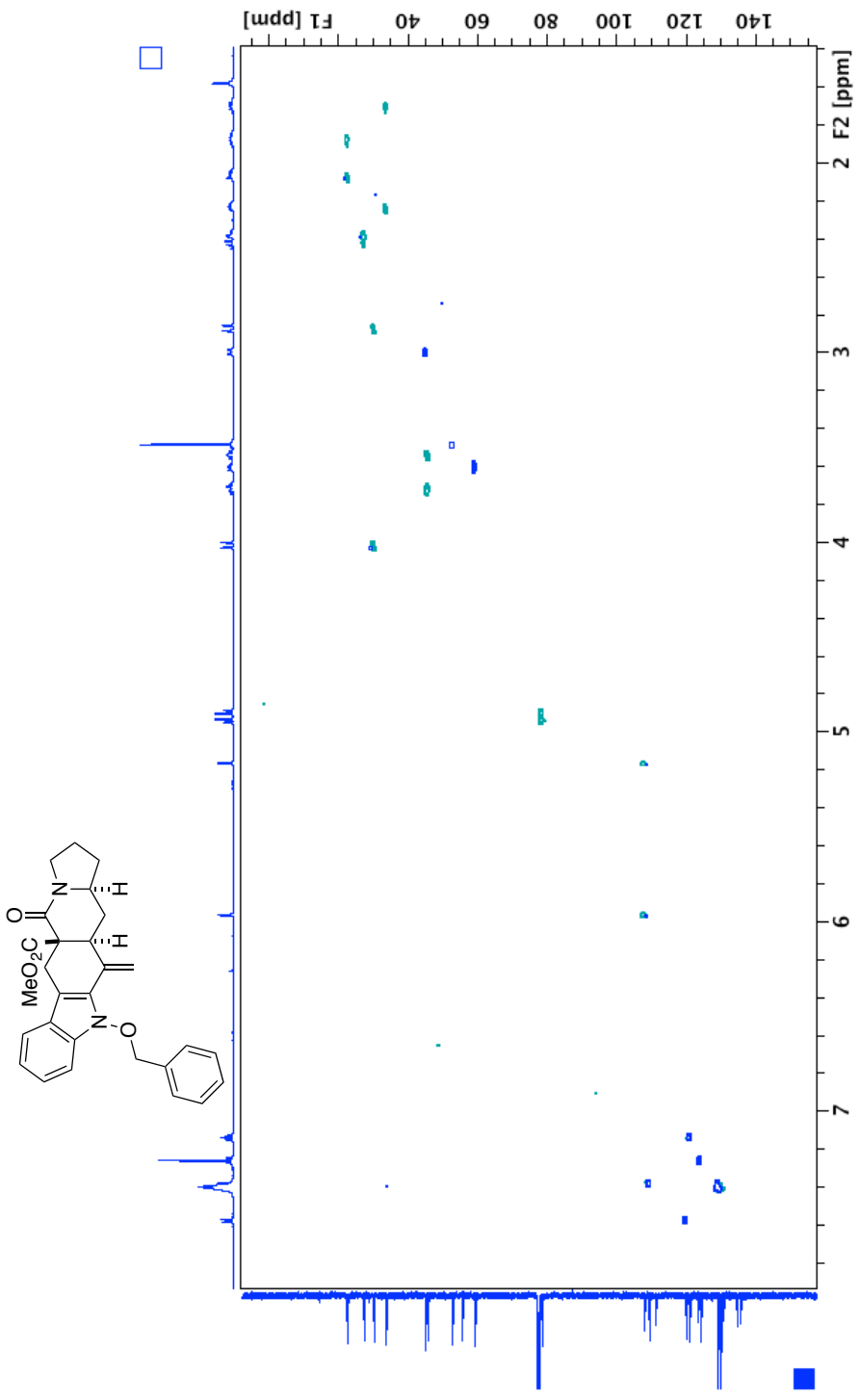
^{1}H NMR of compound **34** (600 MHz, $CDCl_3$)



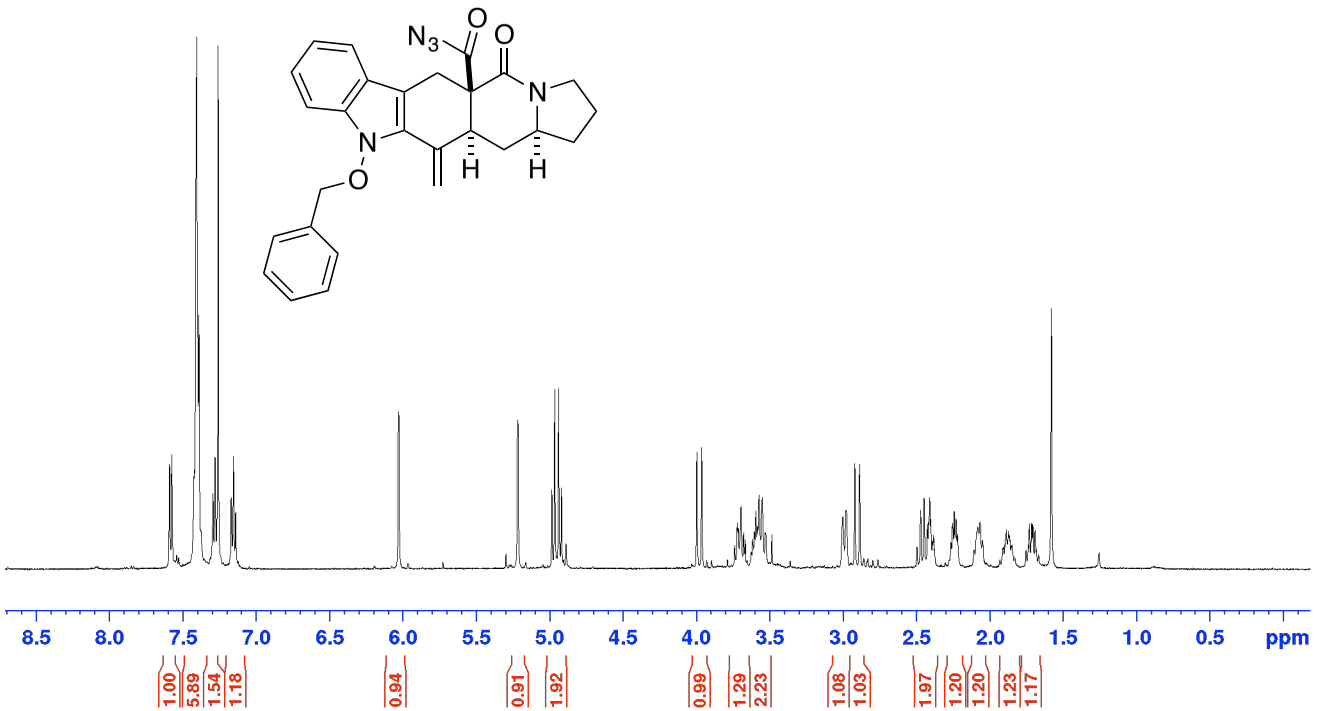
^{13}C NMR of compound **34** (150 MHz, $CDCl_3$)



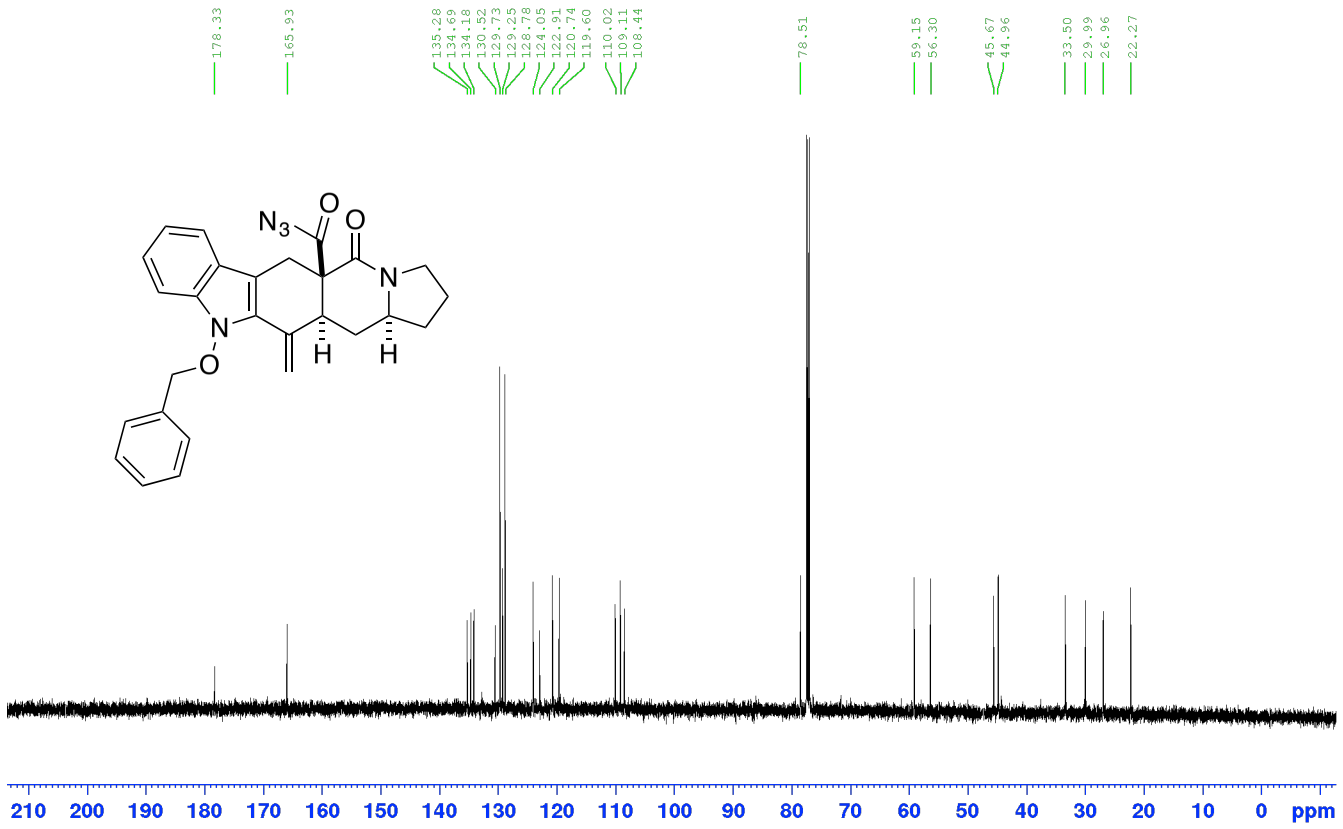
COSY spectrum of compound **34** (600 MHz, $CDCl_3$)



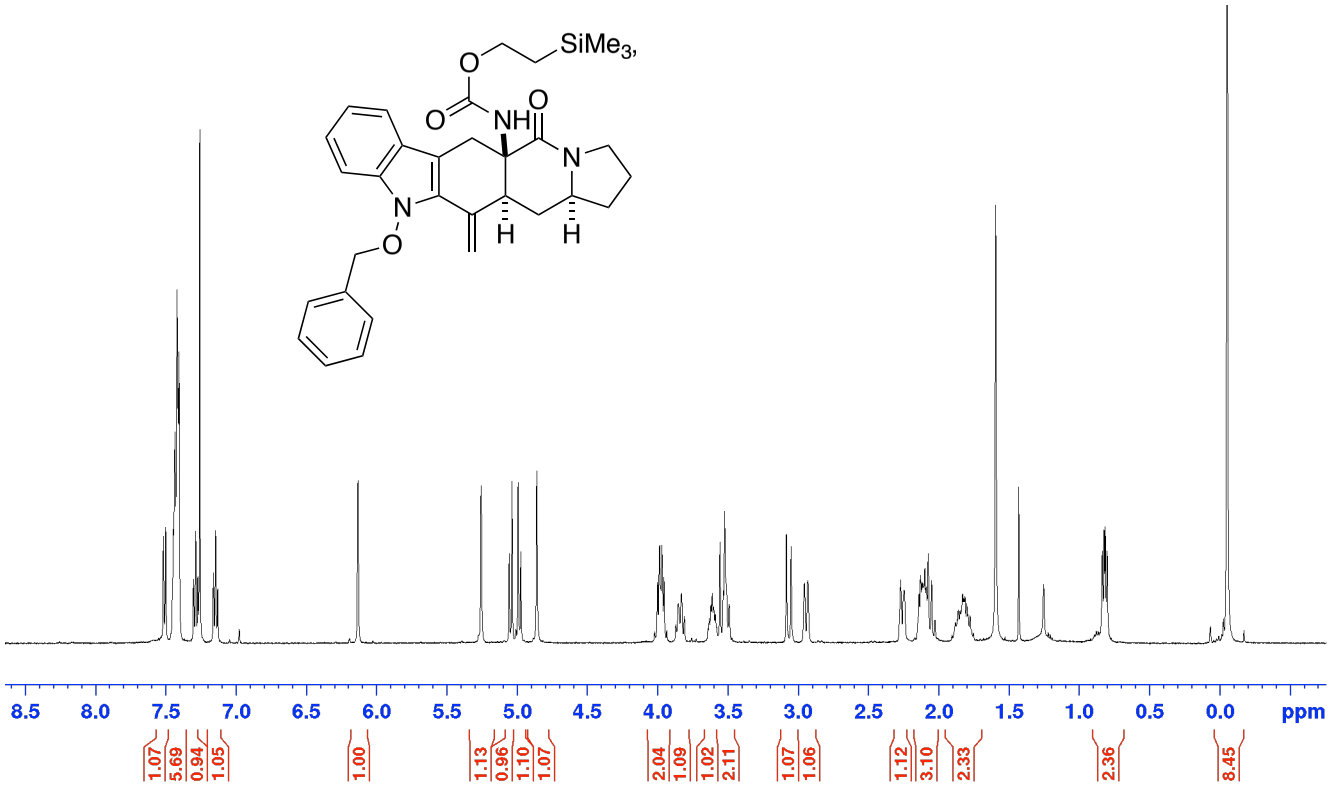
HSQC spectrum of compound **34** (600 MHz, $CDCl_3$)



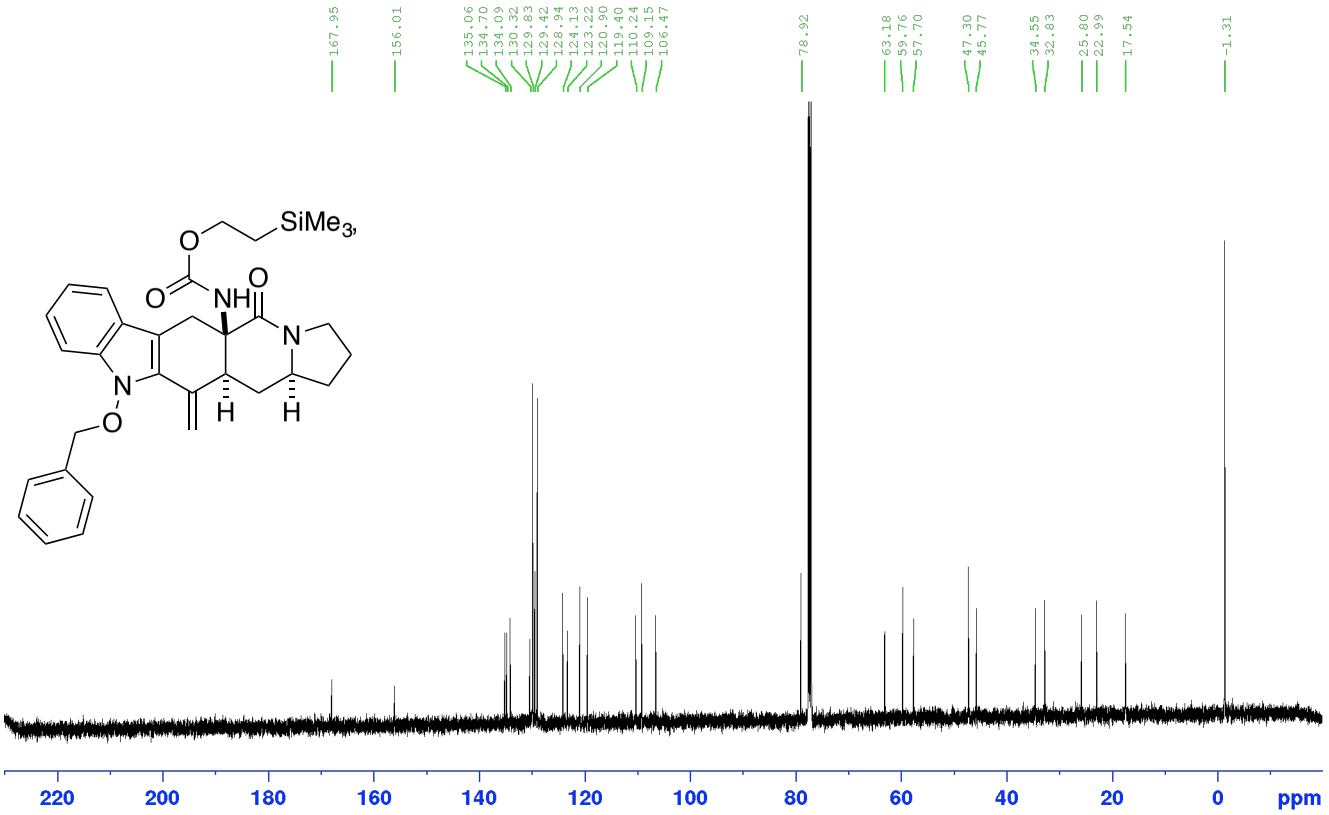
^{1}H NMR of compound **S-1** (500 MHz, $CDCl_3$)



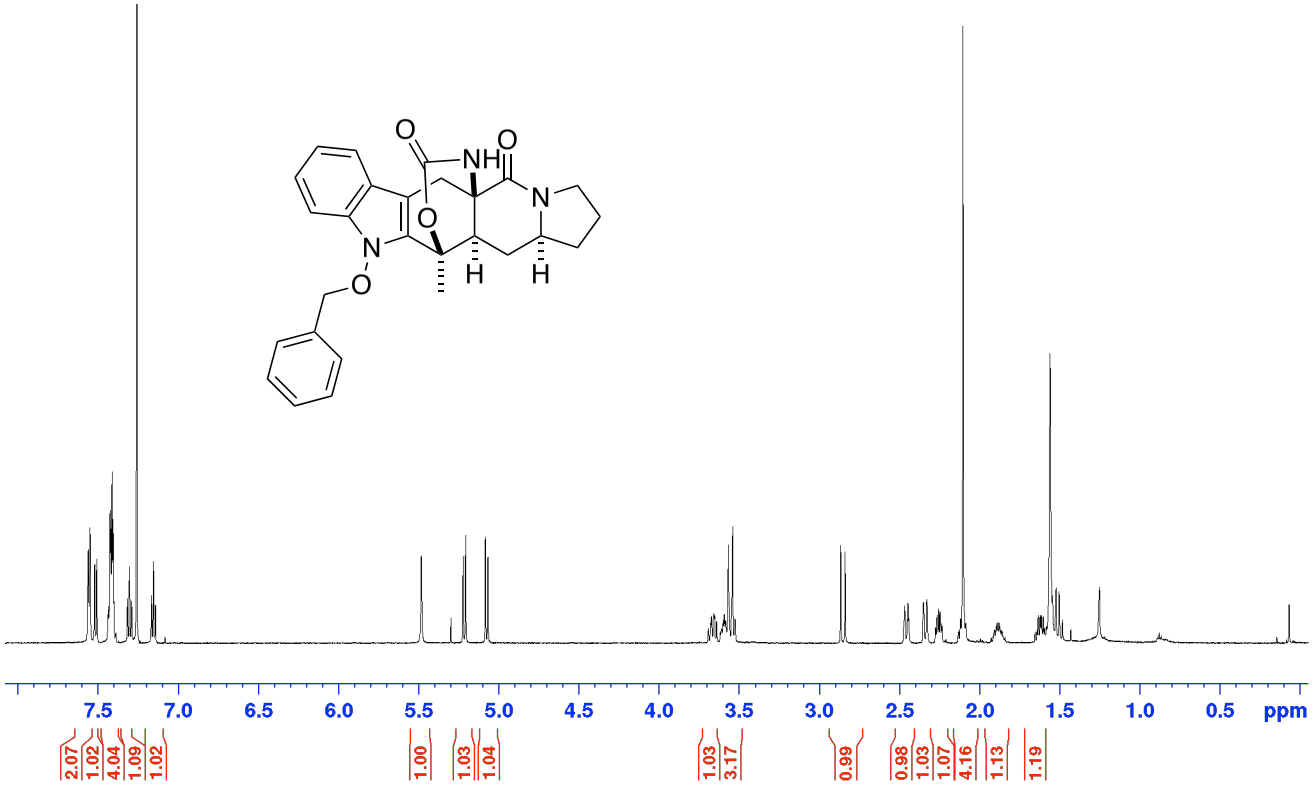
^{13}C NMR of compound **S-1** (125 MHz, $CDCl_3$)



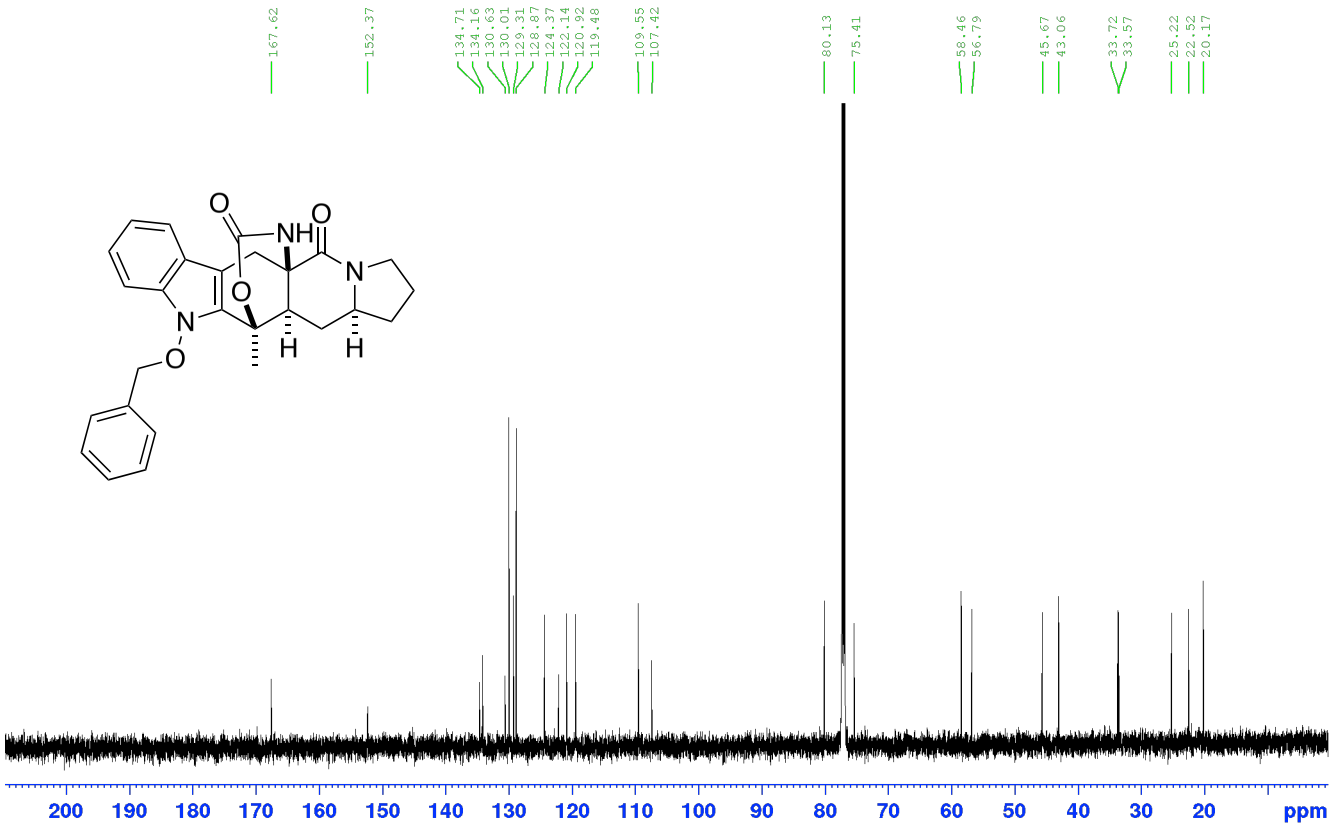
^{1}H NMR of compound **35** (500 MHz, $CDCl_3$)



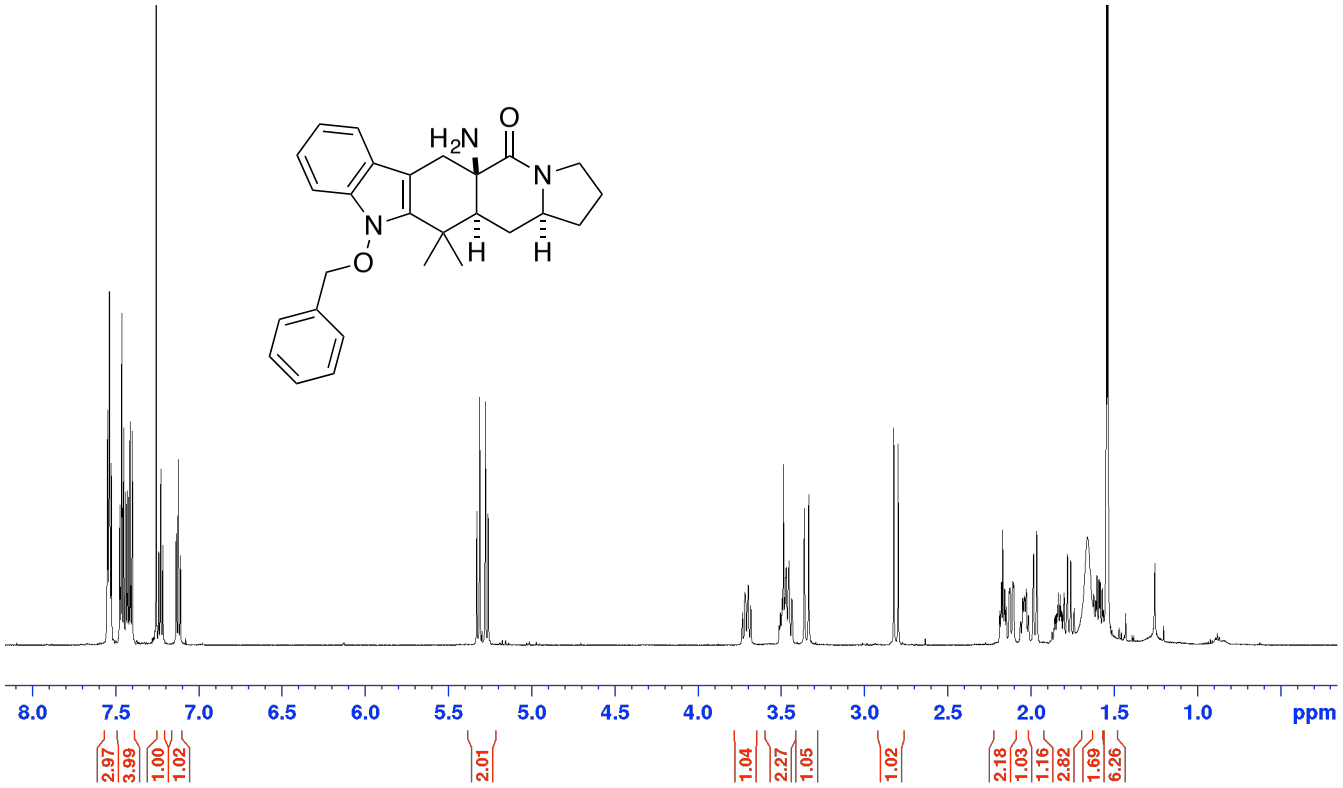
^{13}C NMR of compound **35** (125 MHz, $CDCl_3$)



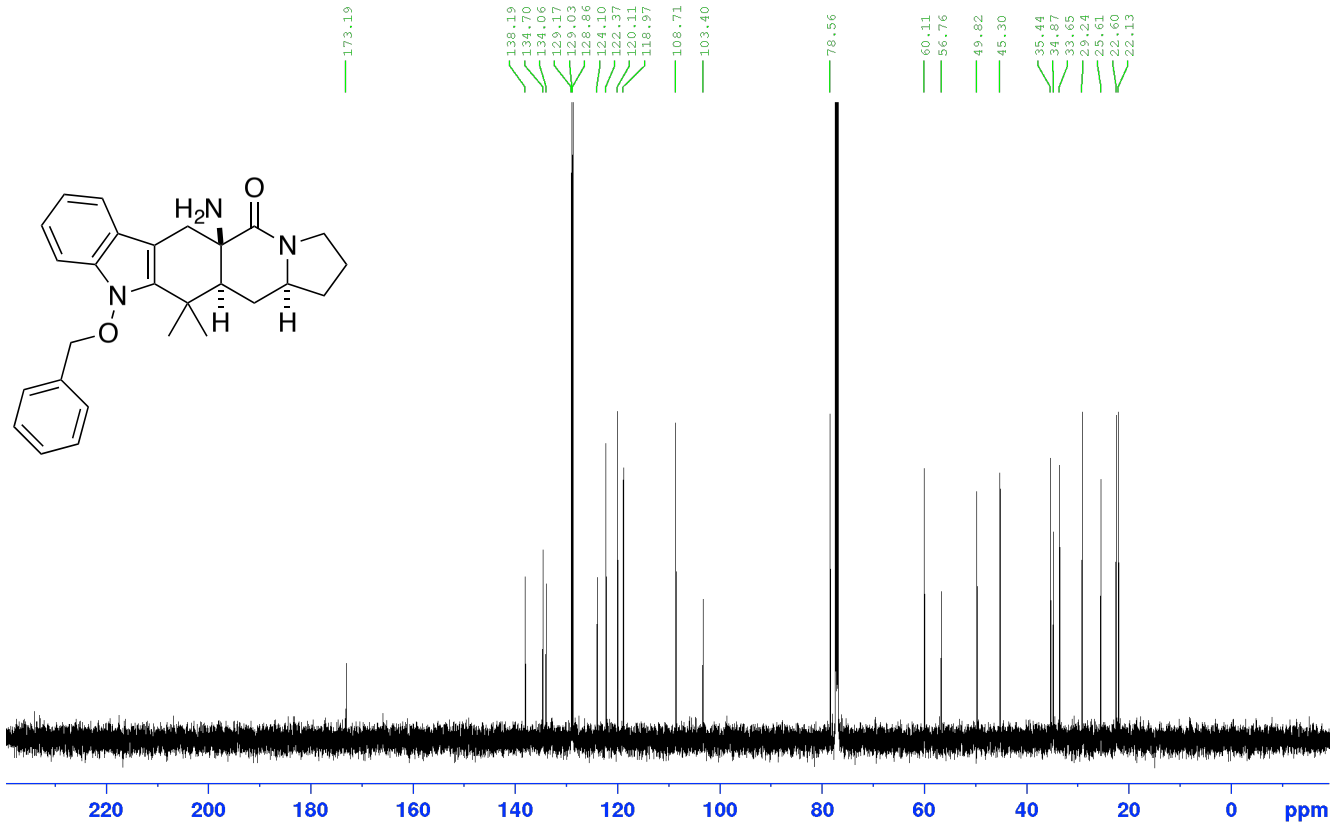
^{1}H NMR of compound **40** (600 MHz, $CDCl_3$)



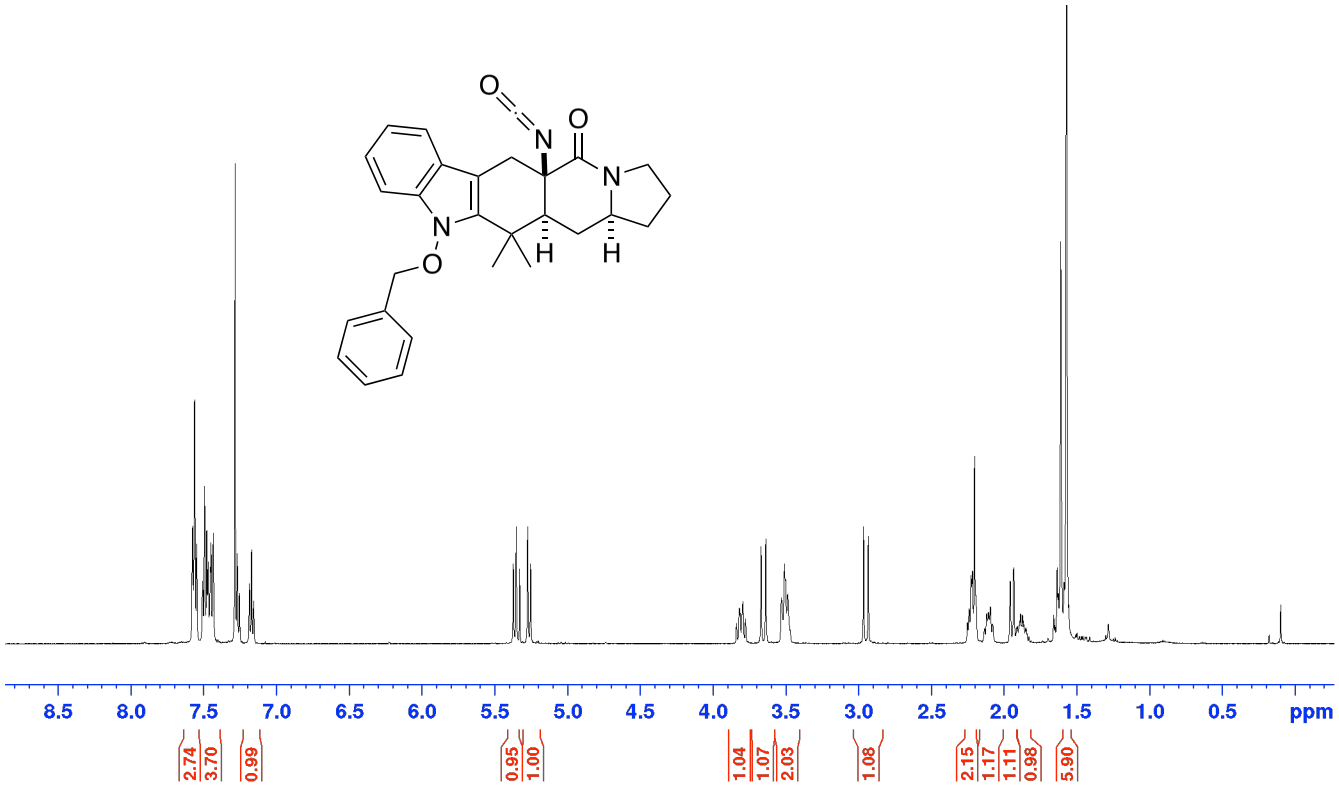
^{13}C NMR of compound **40** (150 MHz, $CDCl_3$)



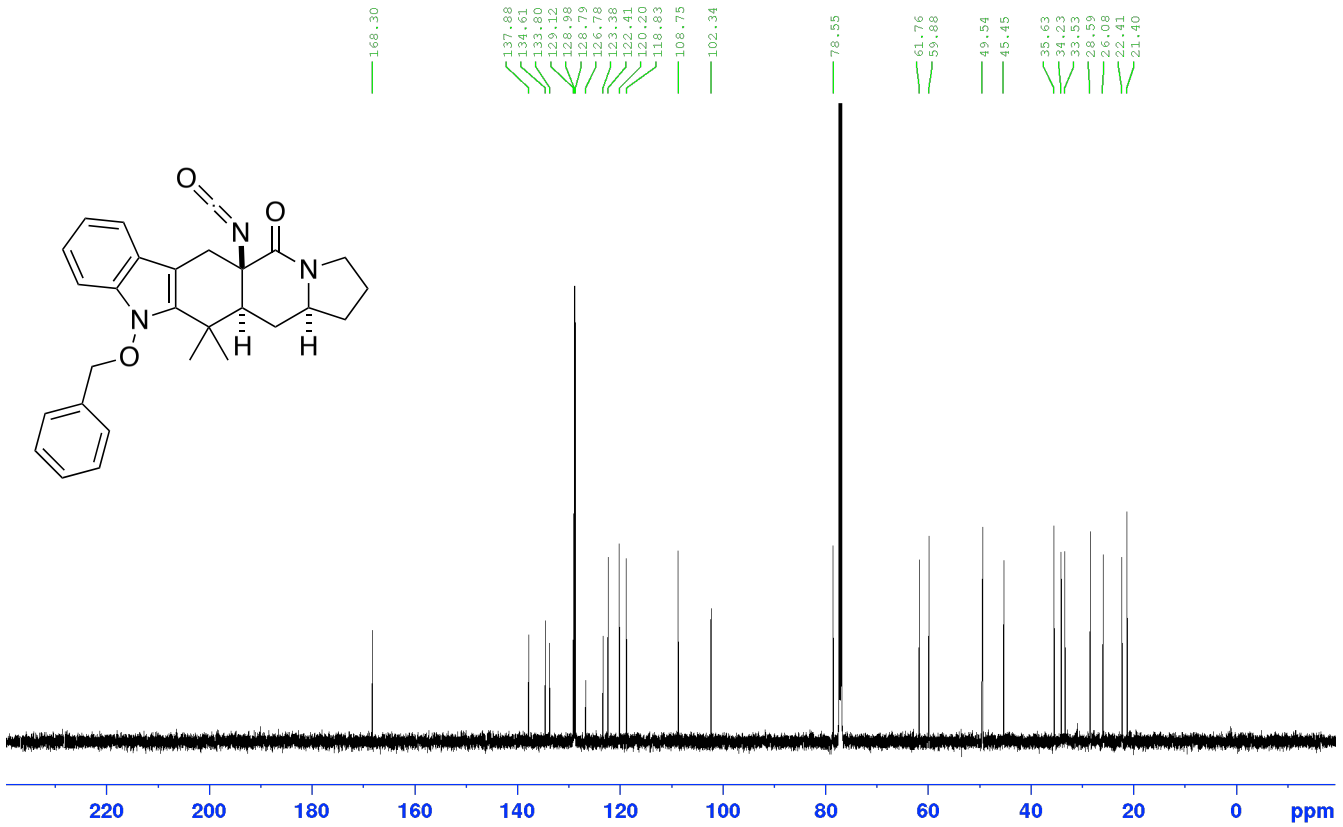
^{1}H NMR of compound **42** (600 MHz, $CDCl_3$)



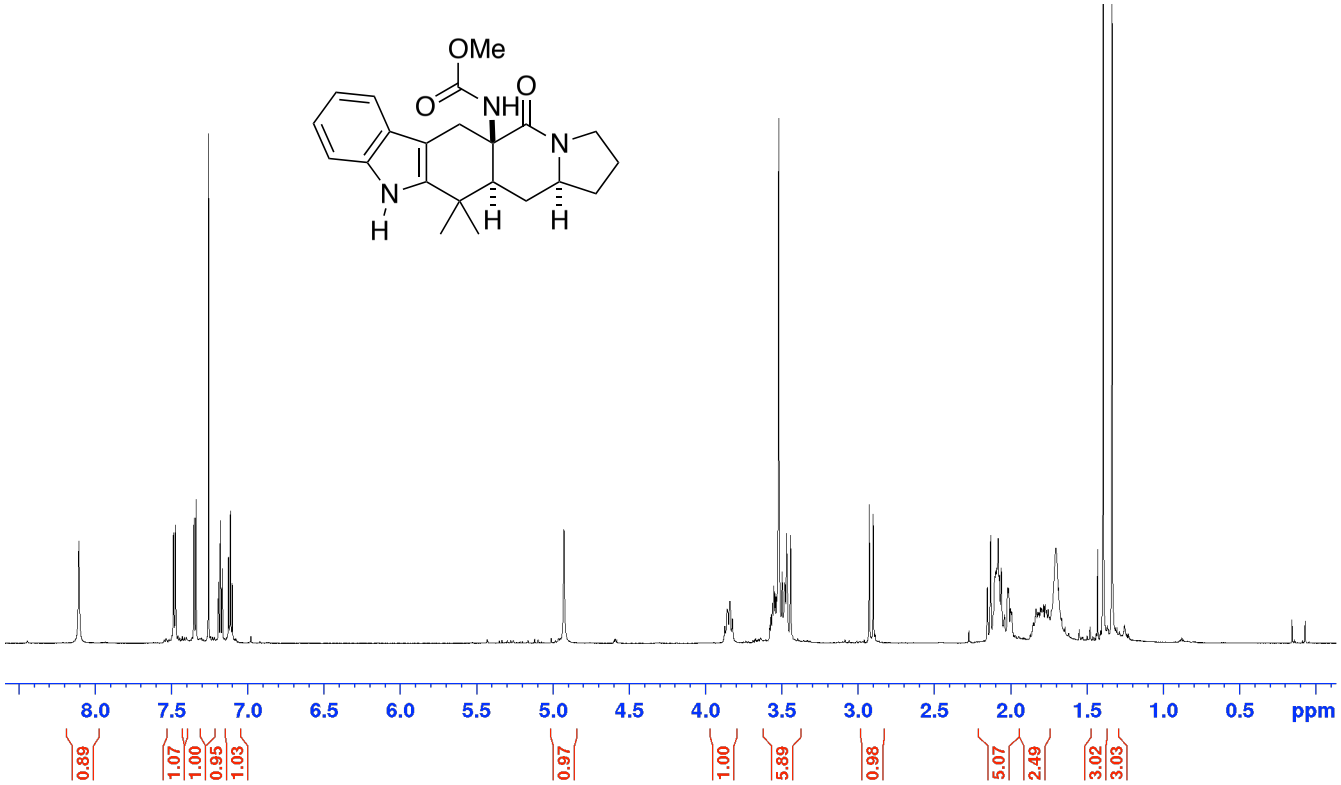
^{13}C NMR of compound **42** (150 MHz, $CDCl_3$)



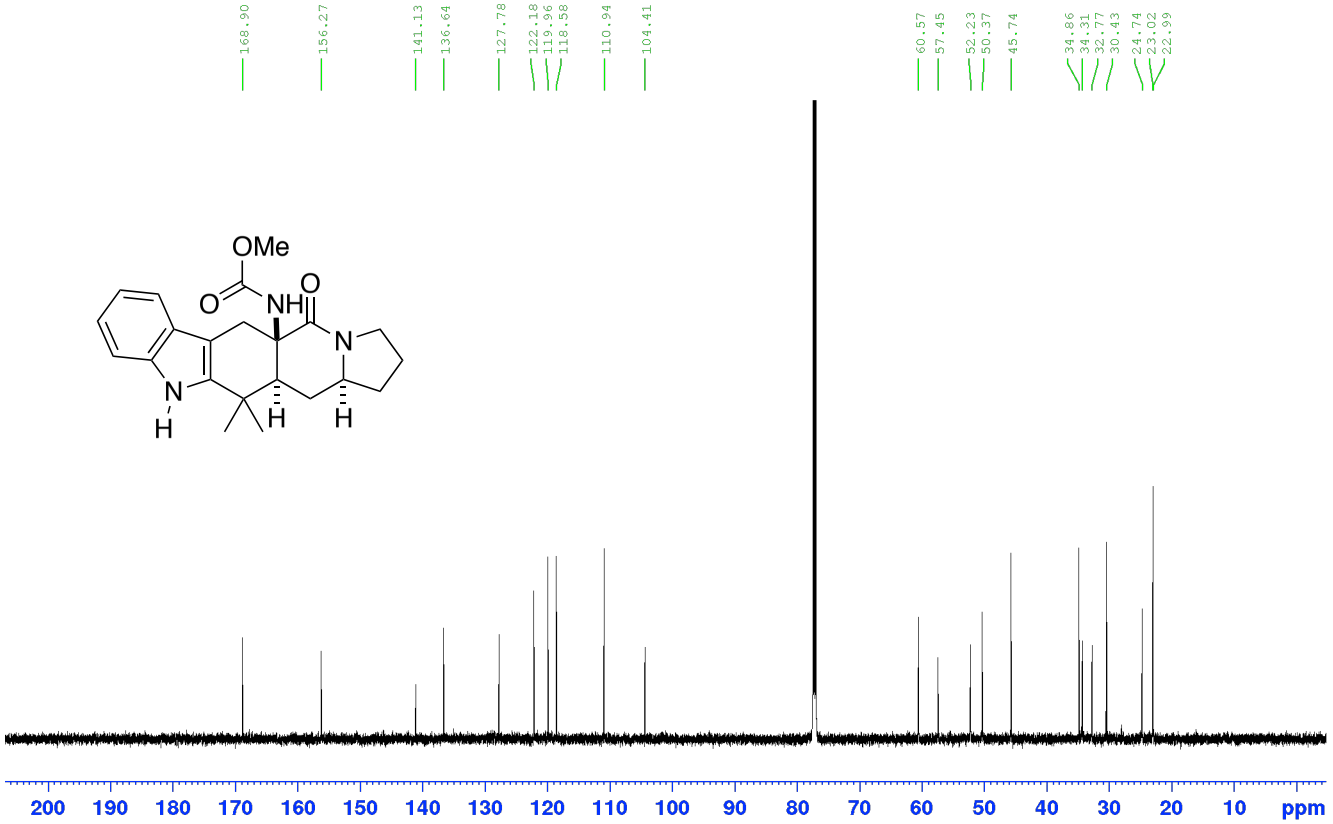
^{1}H NMR of compound **43** (500 MHz, $CDCl_3$)



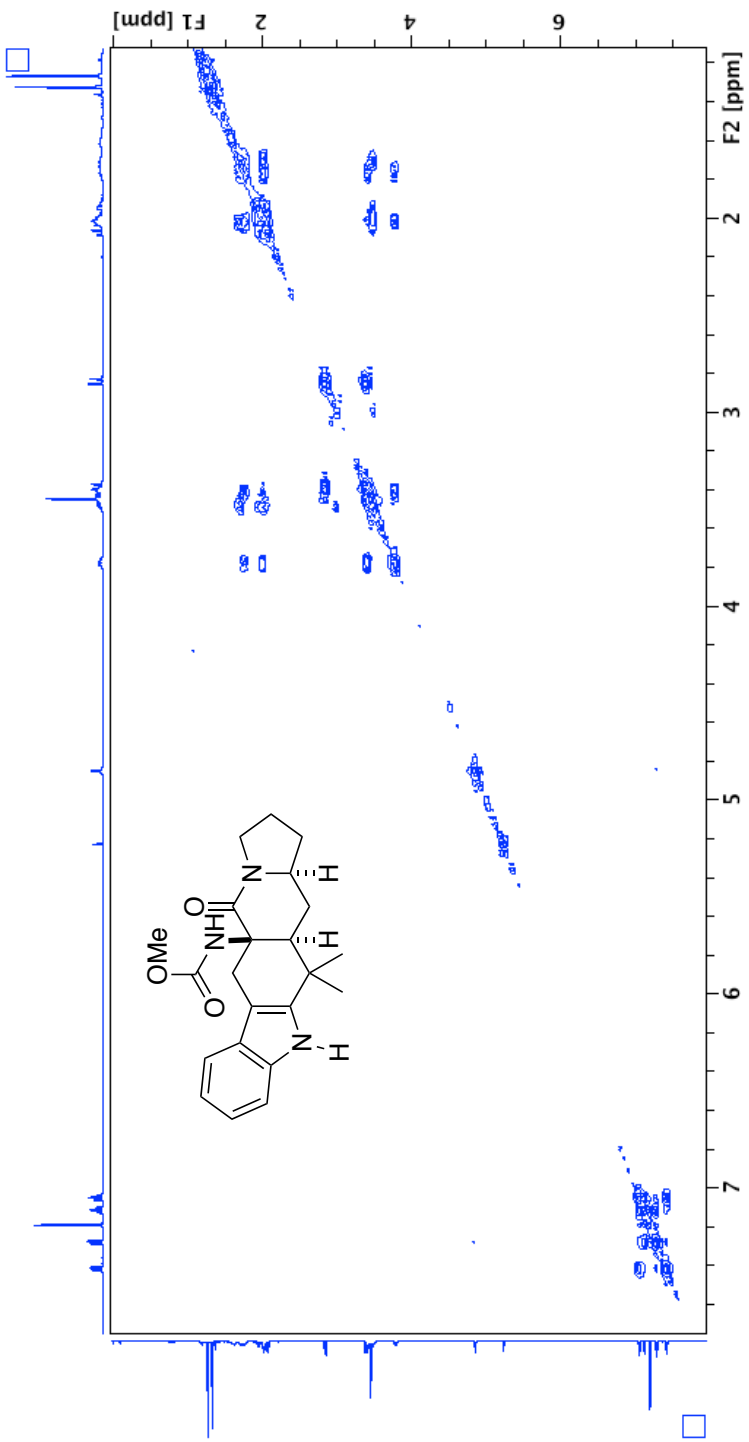
^{13}C NMR of compound **43** (150 MHz, $CDCl_3$)



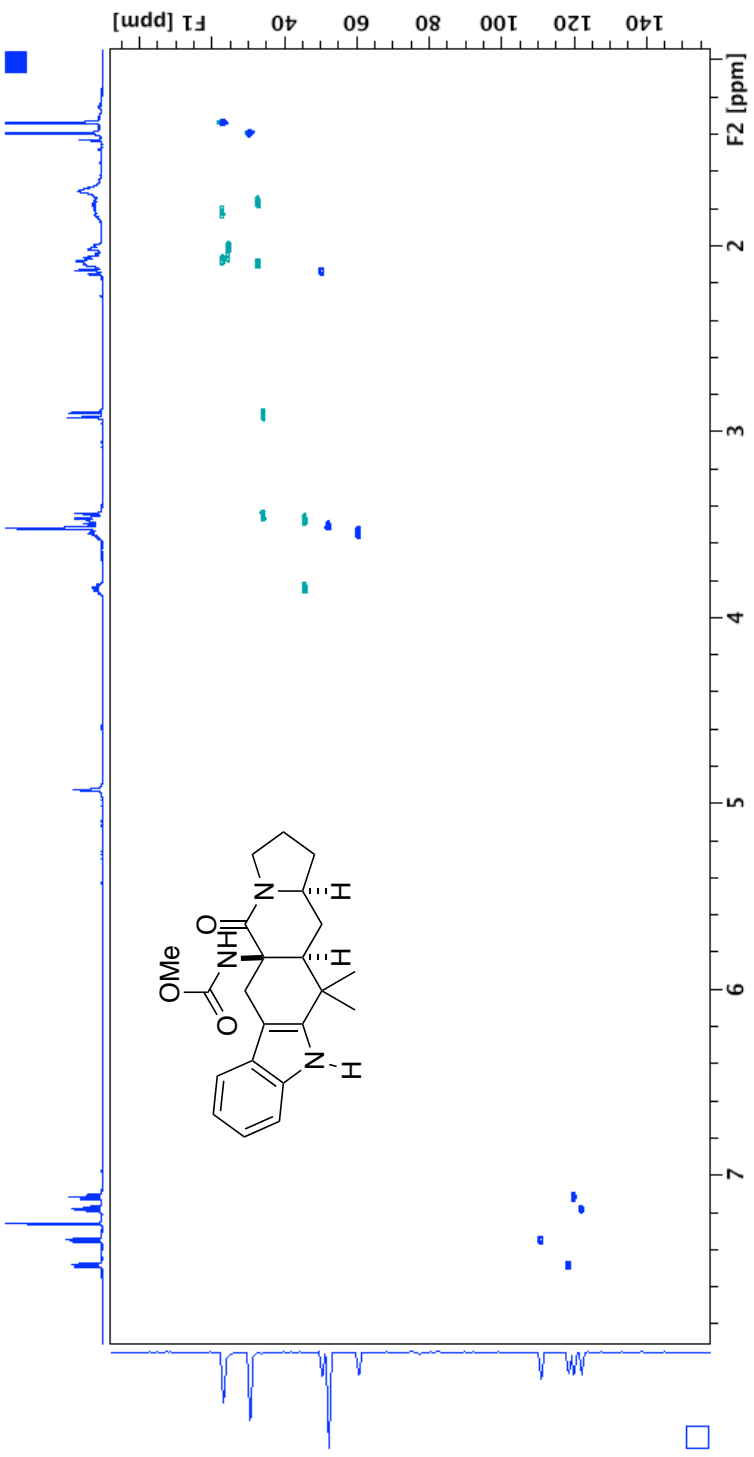
^{1}H NMR of compound **44** (600 MHz, $CDCl_3$)



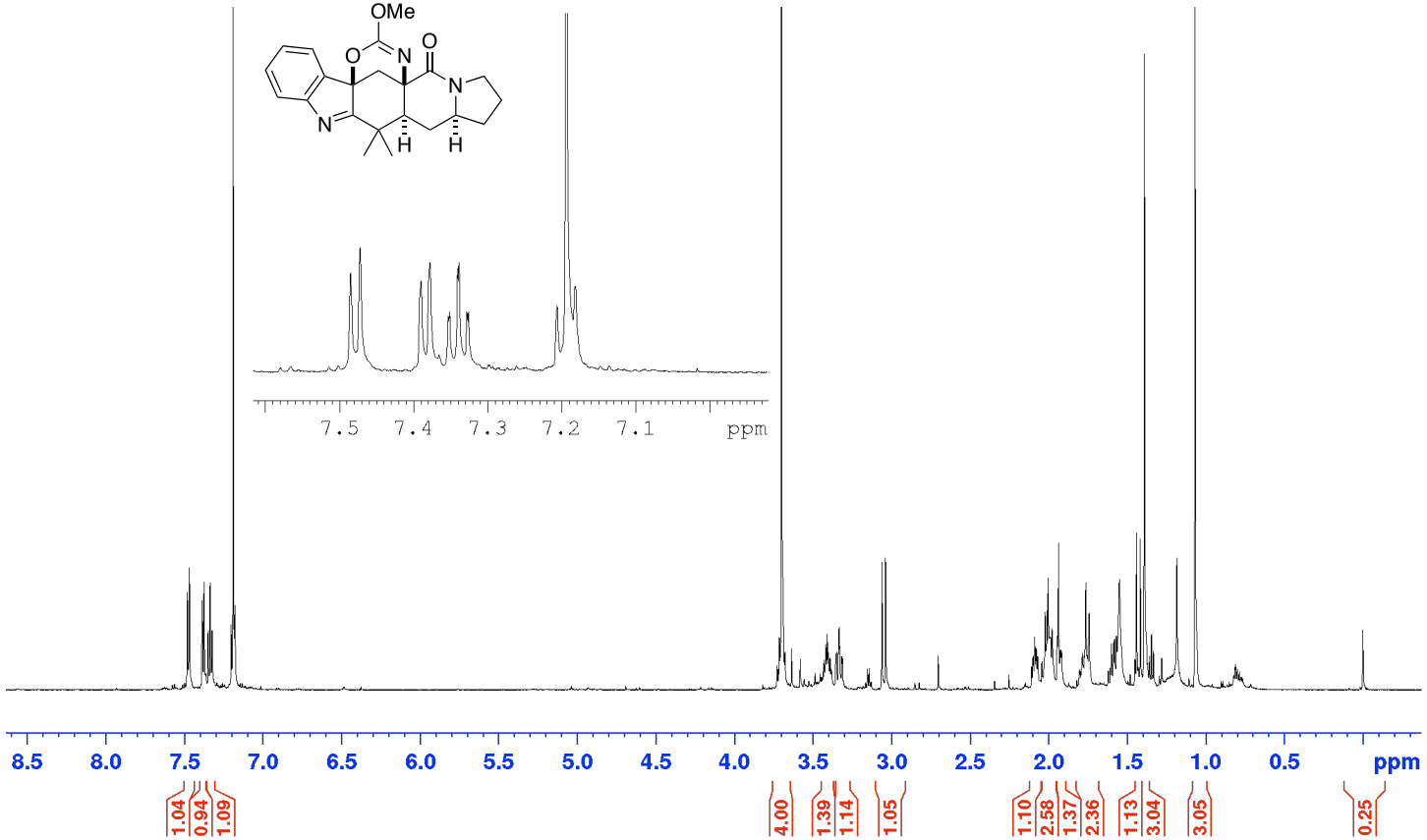
^{13}C NMR of compound **44** (150 MHz, $CDCl_3$)



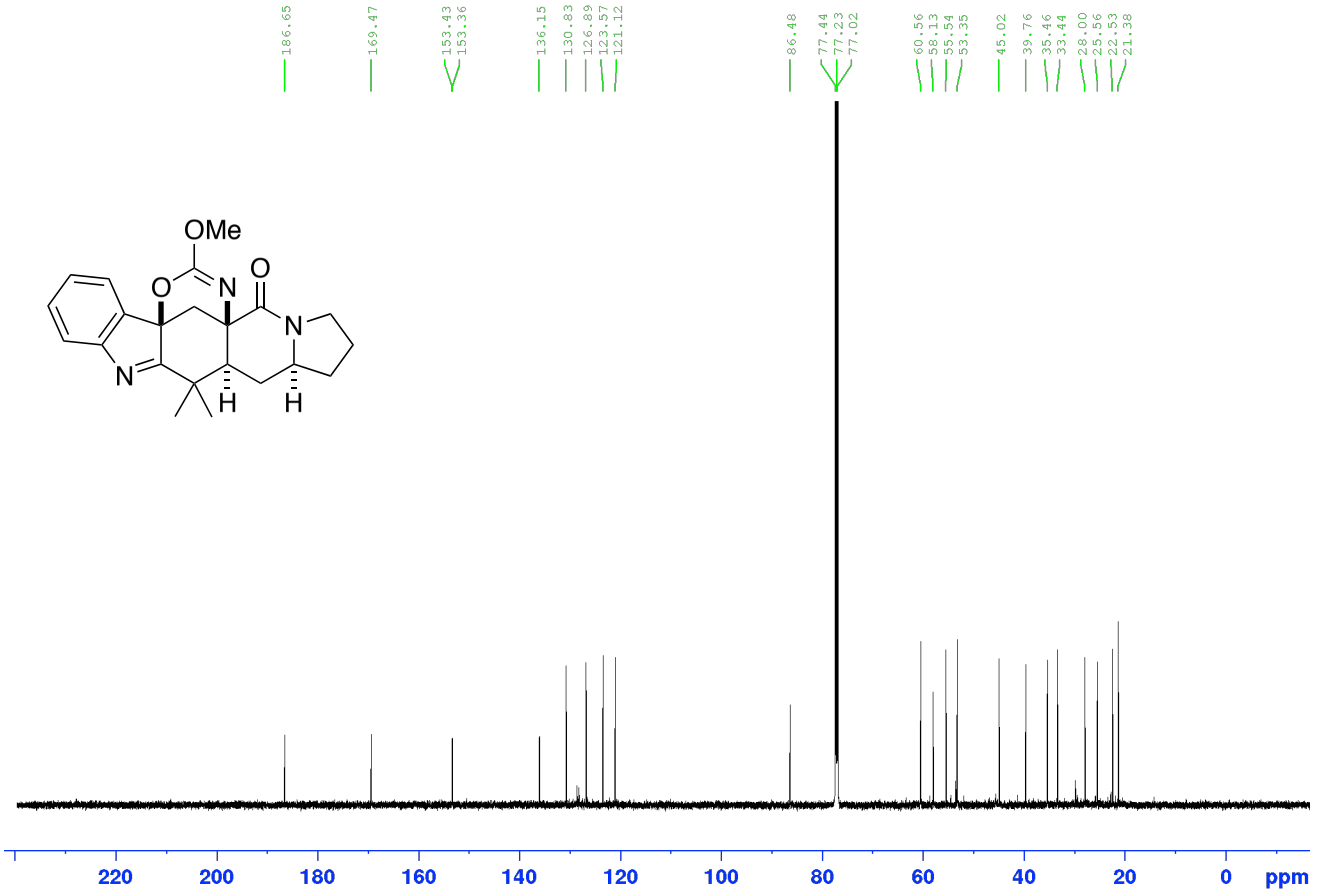
COSY spectrum of compound **44** (600 MHz, $CDCl_3$)



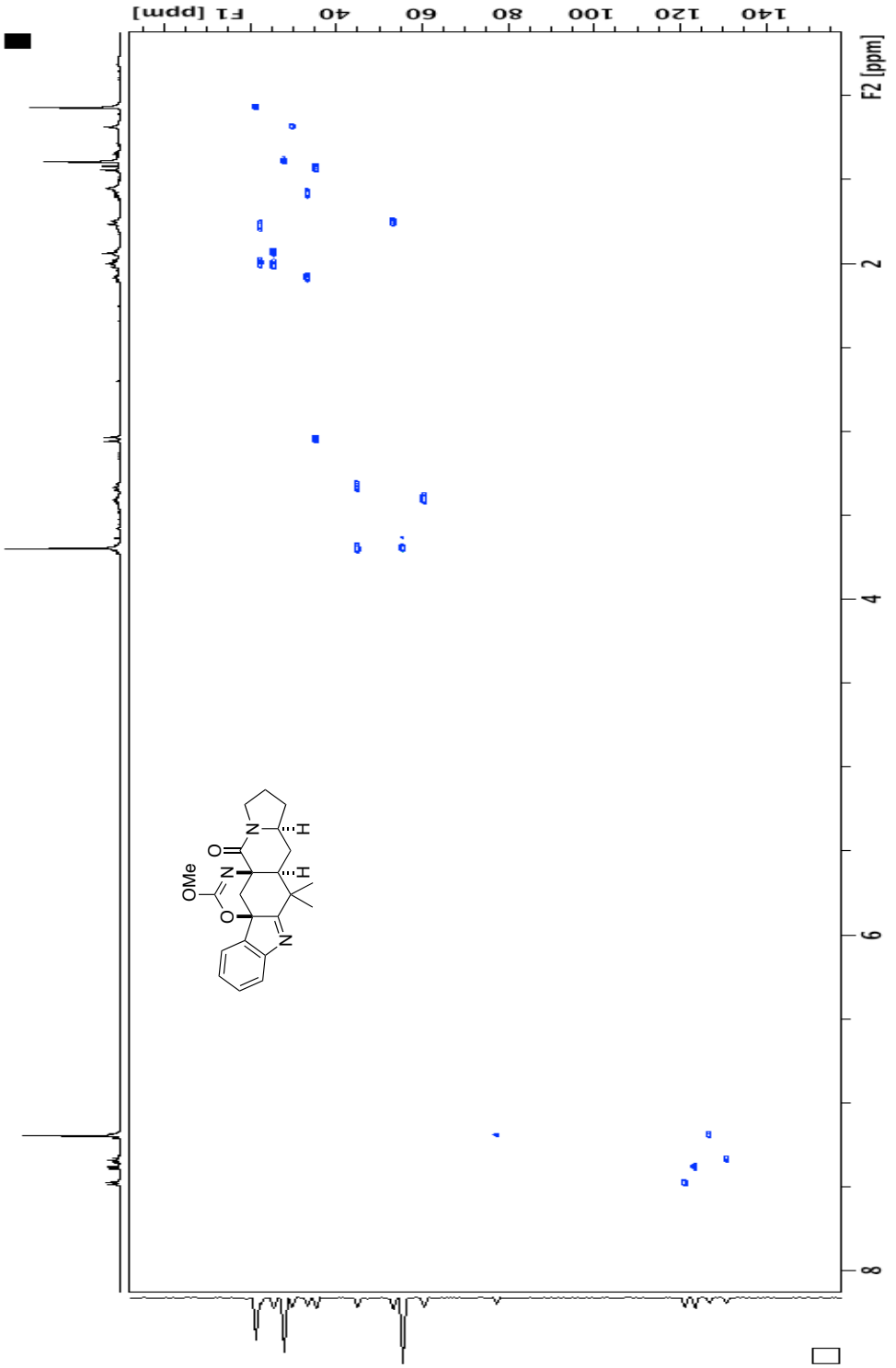
HSQC spectrum of compound **44** (600 MHz, $CDCl_3$)



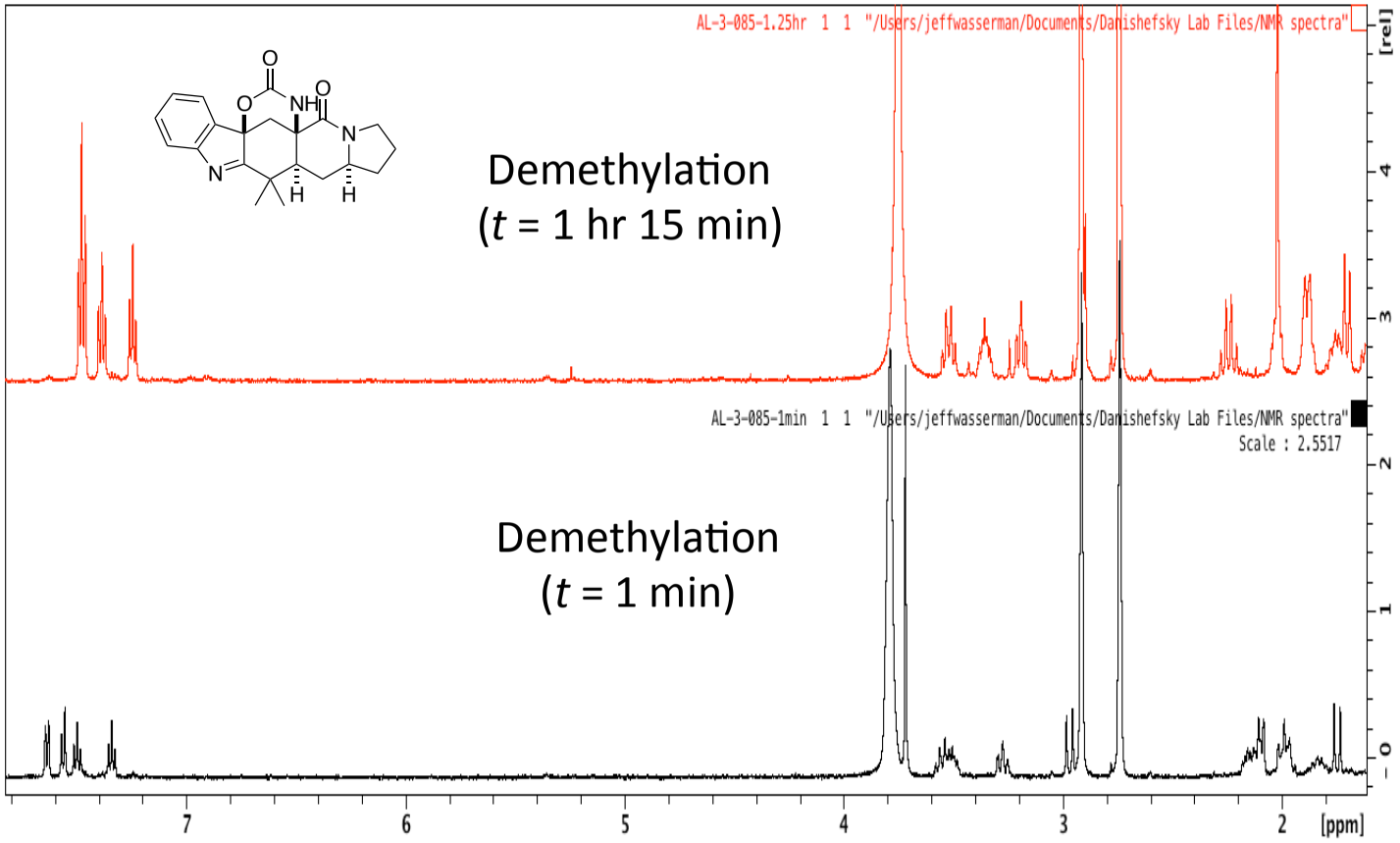
^{1}H NMR of compound **56** (600 MHz, $CDCl_3$)



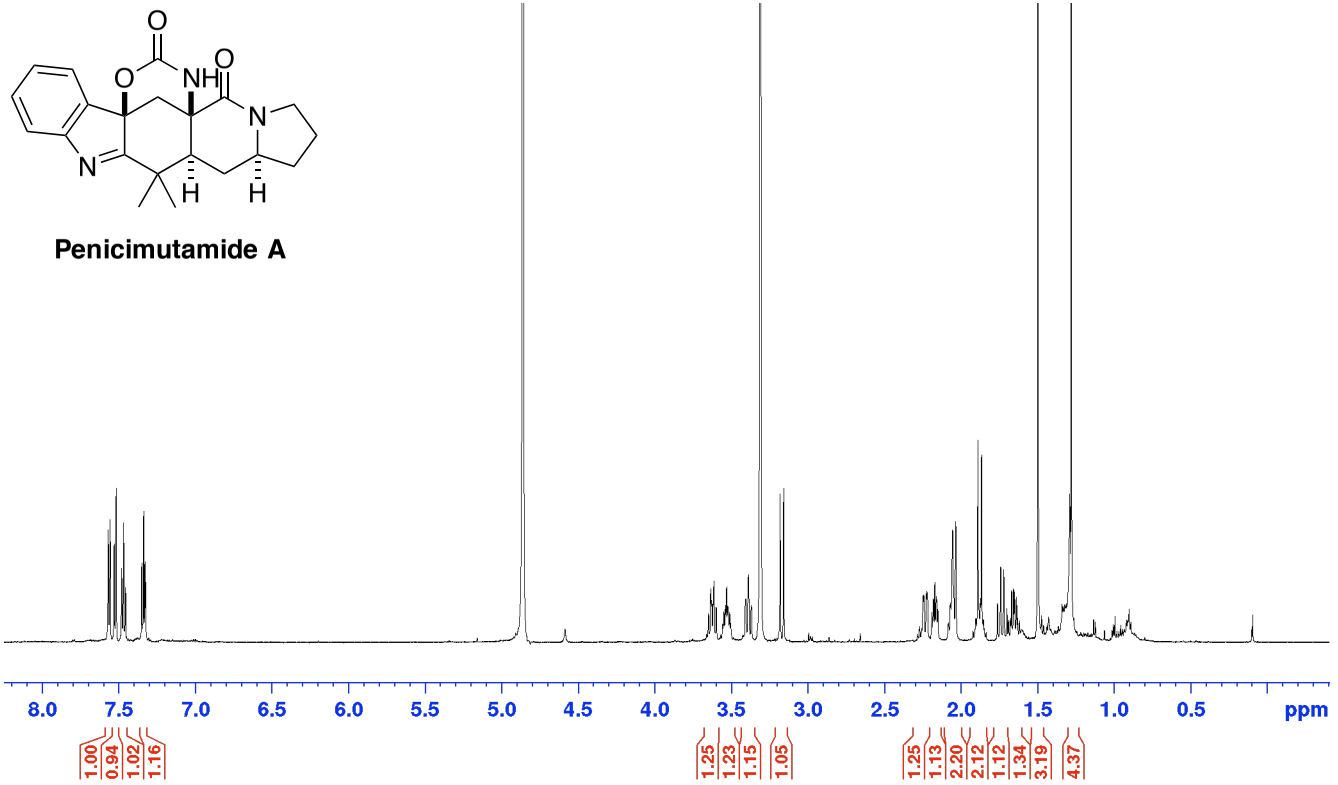
^{31}C NMR of compound **56** (150 MHz, $CDCl_3$)



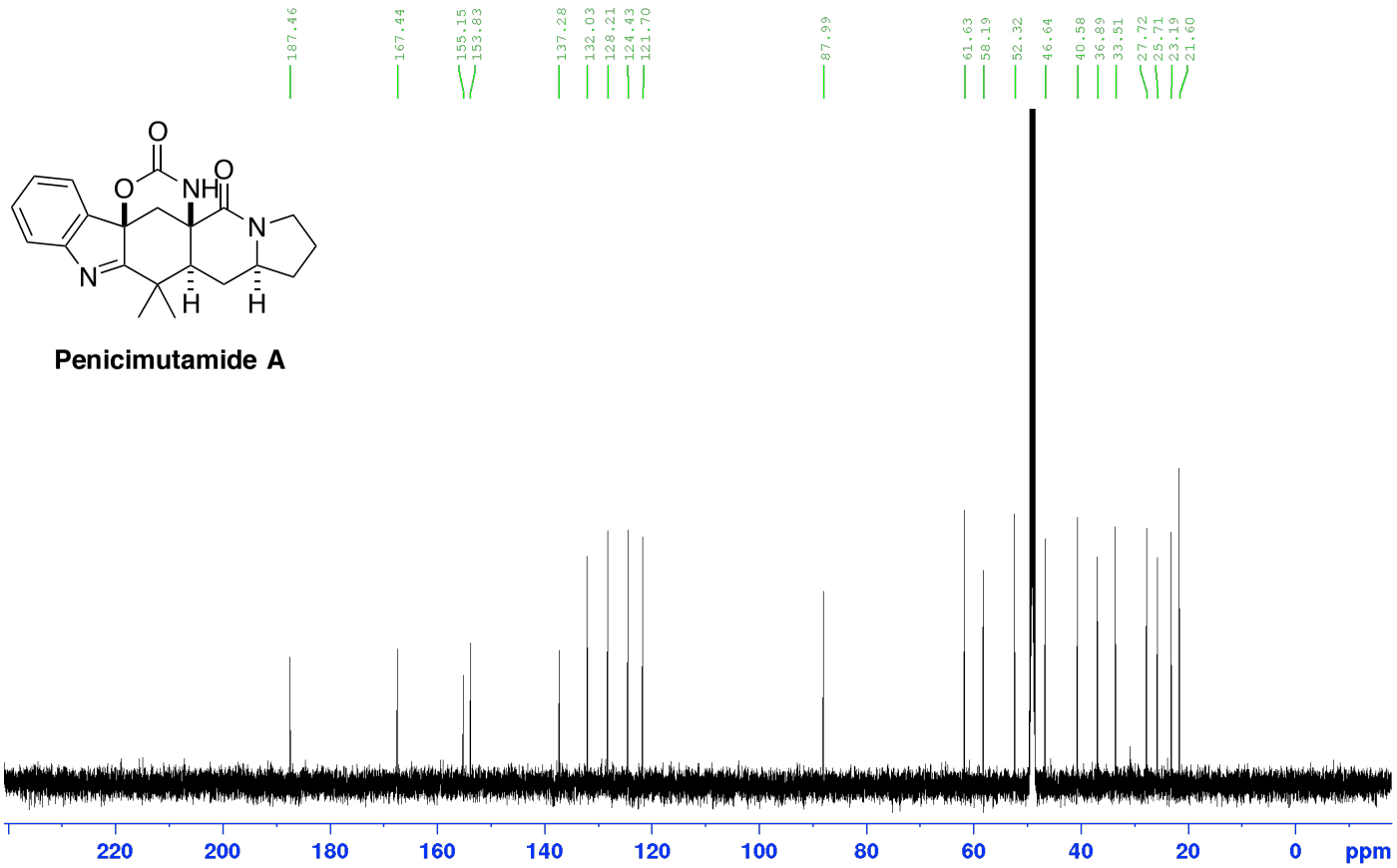
HSQC spectrum of **56** ($CDCl_3$, 600 MHz)



^{1}H NMR analysis of nucleophilic demethylation reaction of **56** to form **2** (500 MHz, DMF-d_7)
-Test reaction run on a 3 mg scale in DMF-d_7. Analysis of the reaction mixture indicates quantitative conversion to penicimutamide A (**2**) at (t = 1 hr 15 min)

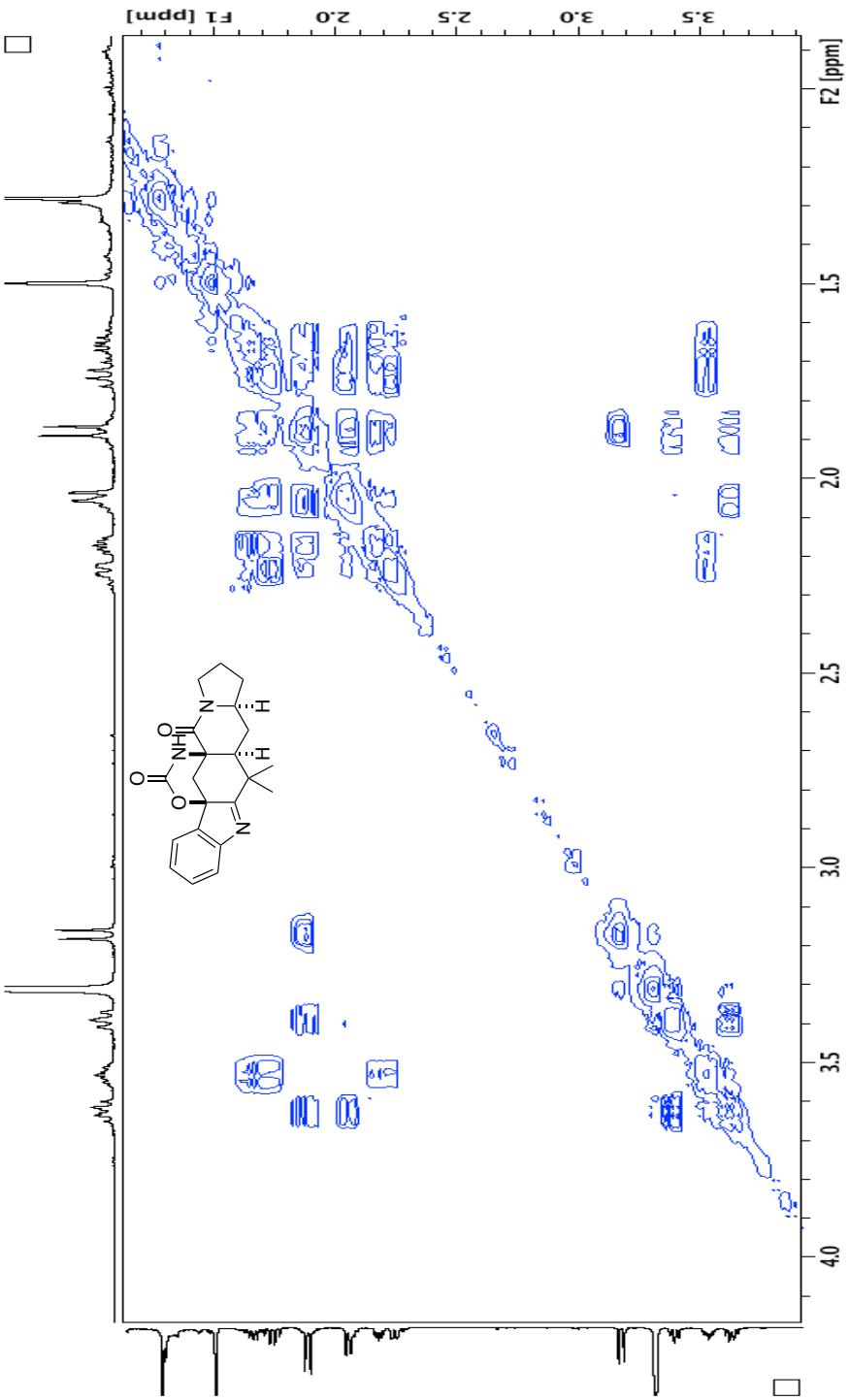


^{1}H NMR of compound **2** (600 MHz, MeOD)

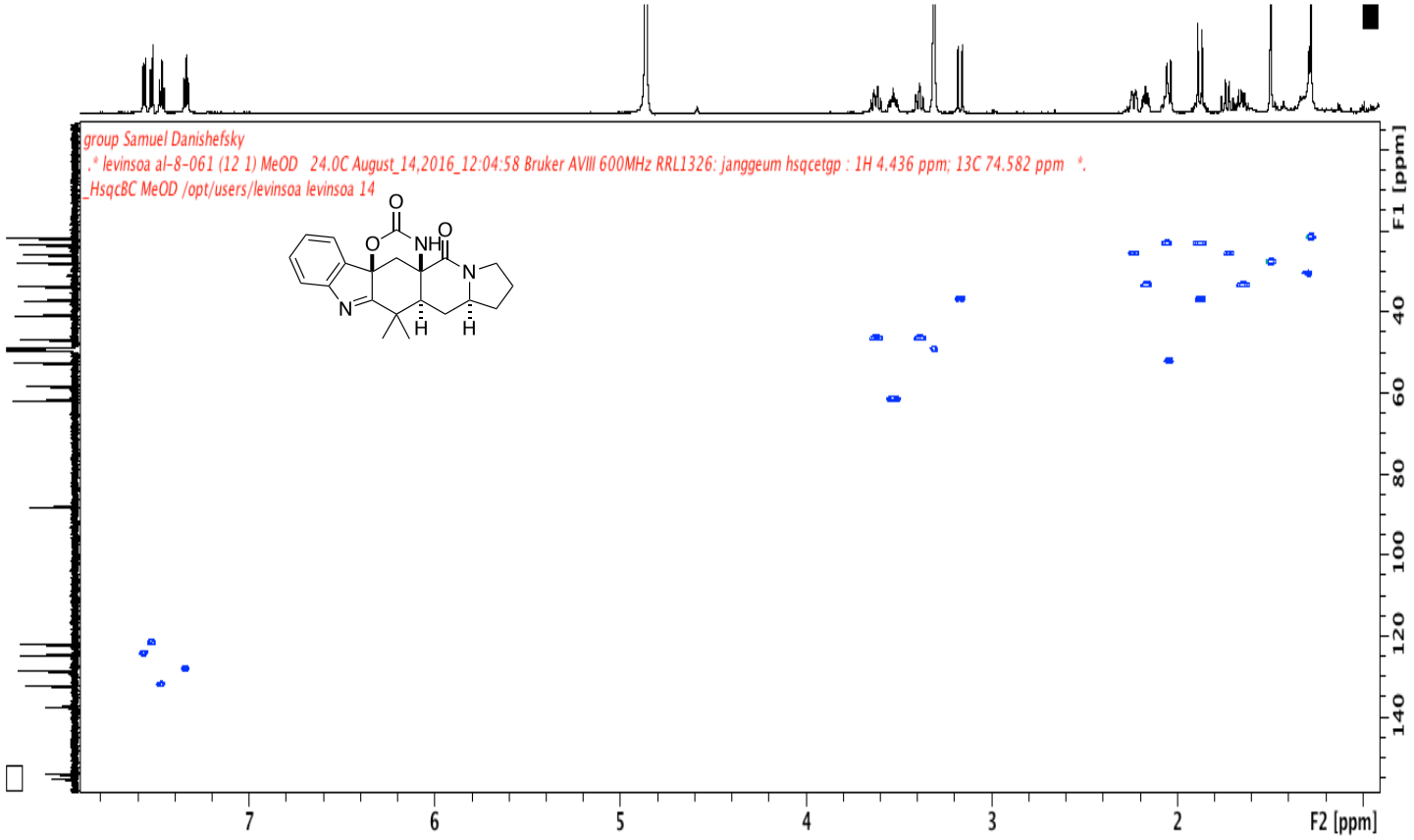


Penicimutamide A

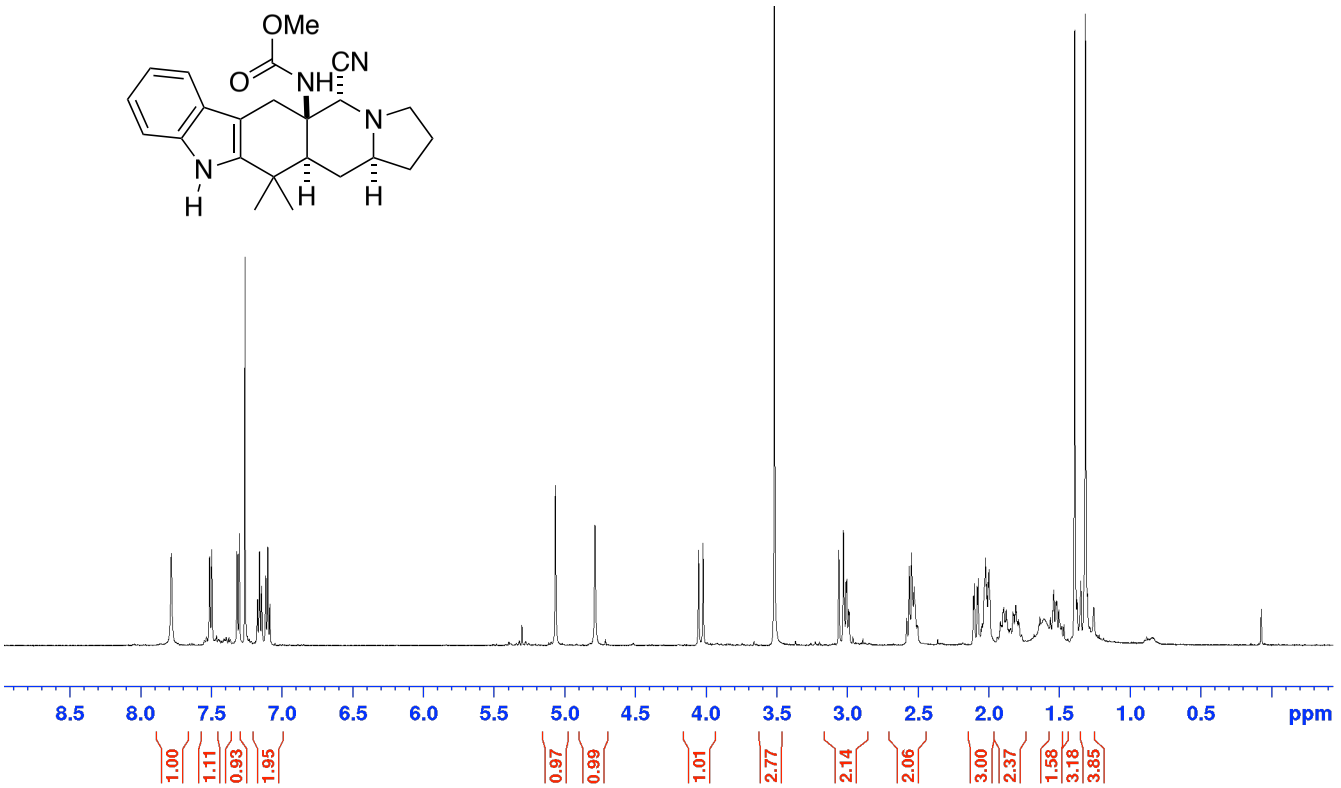
^{13}C NMR of compound **2** (150 MHz, MeOD)



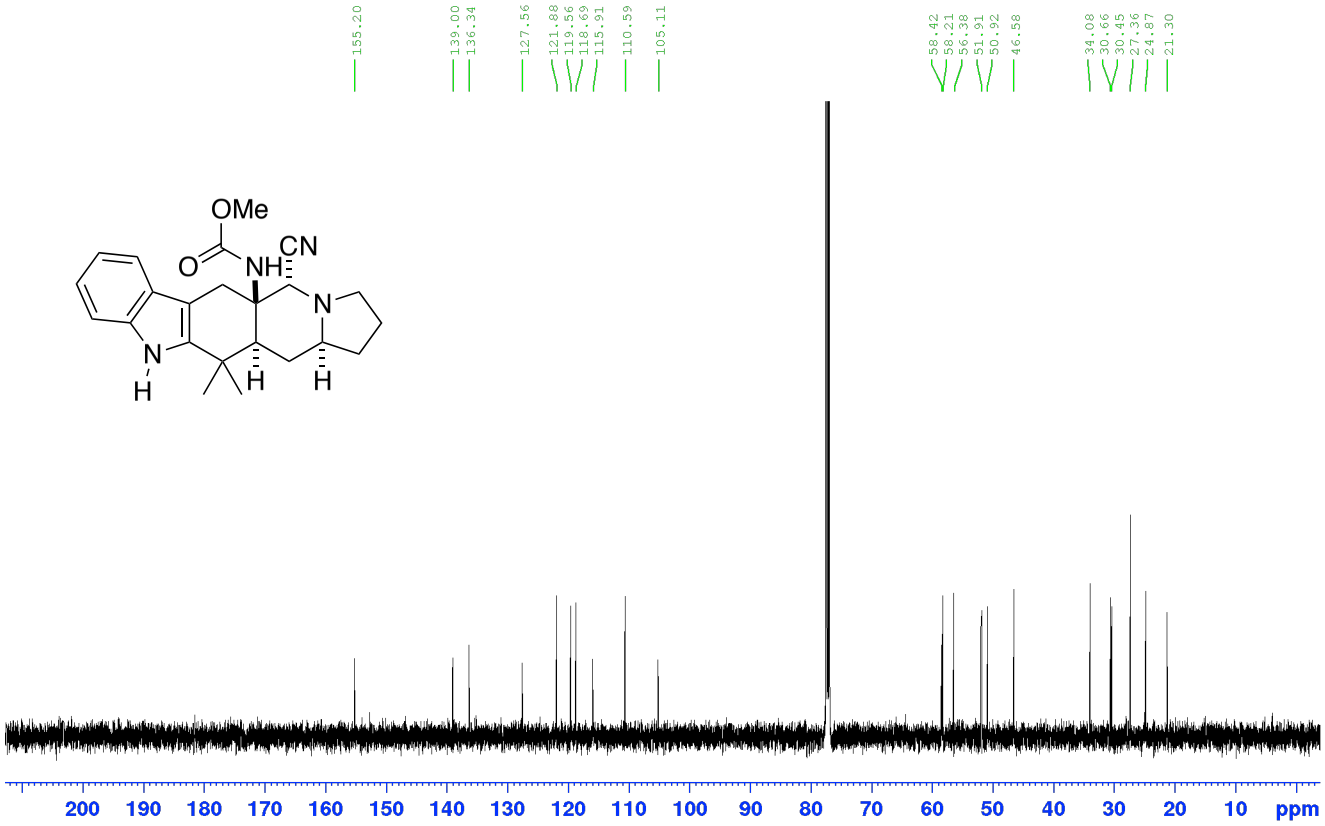
COSY of compound **2** (600 MHz, MeOD)



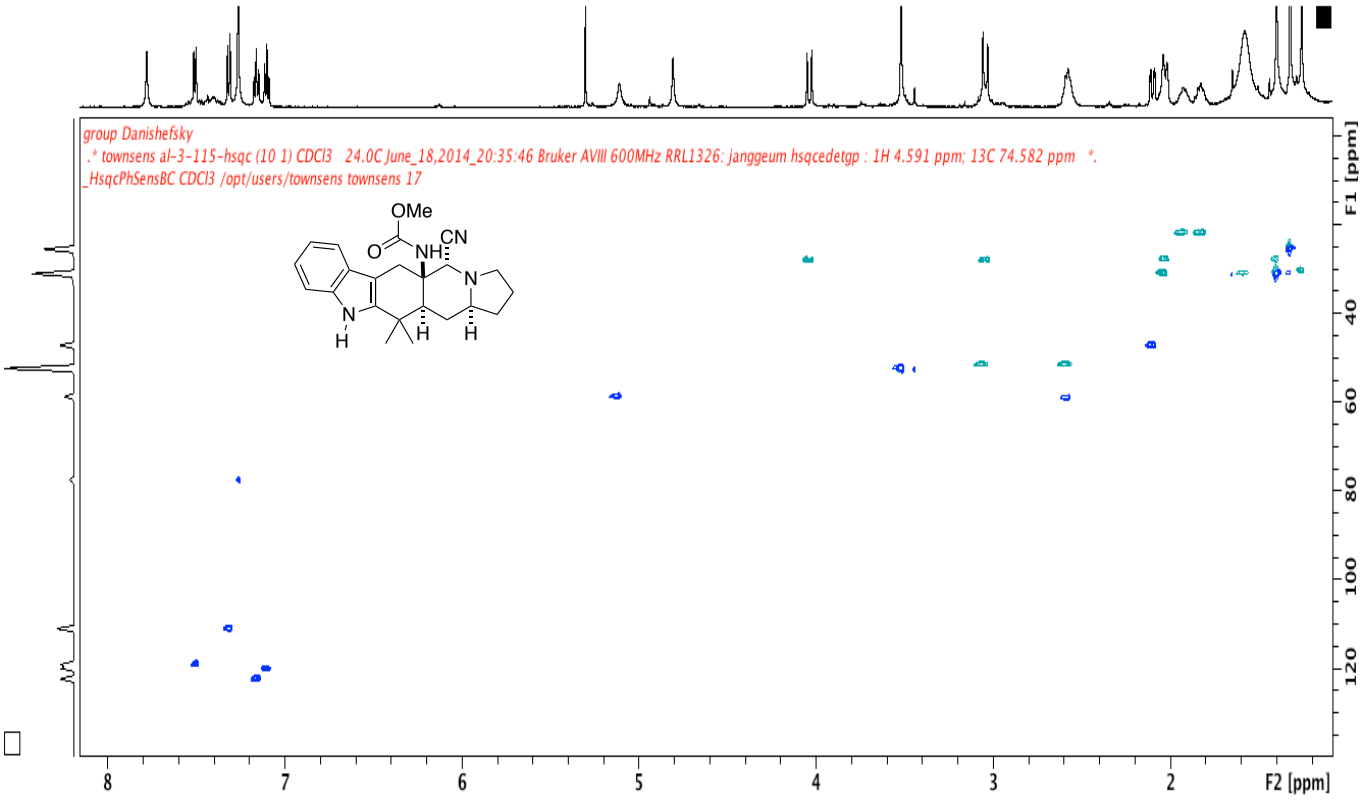
HSQC of compound **2** (600 MHz, MeOD)



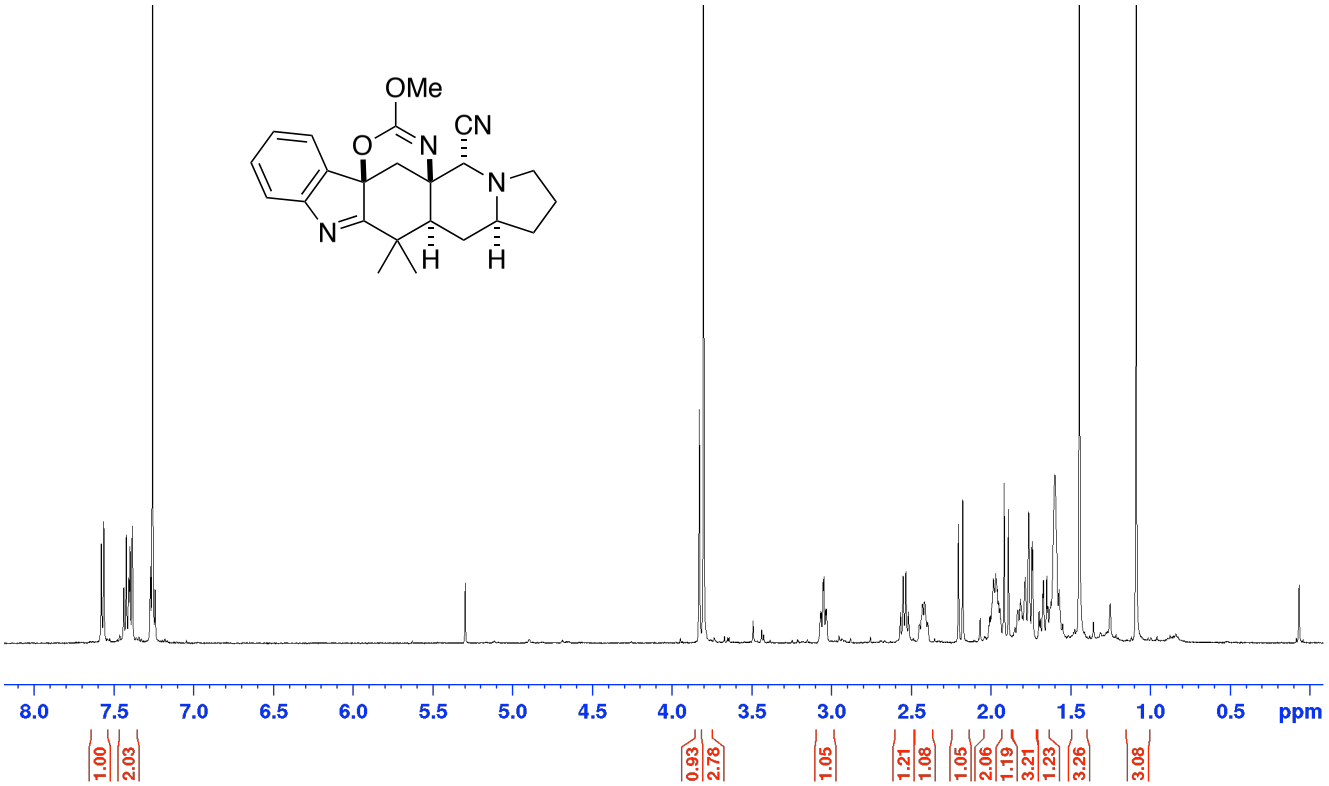
^{1}H NMR of compound **58** (500 MHz, $CDCl_3$)



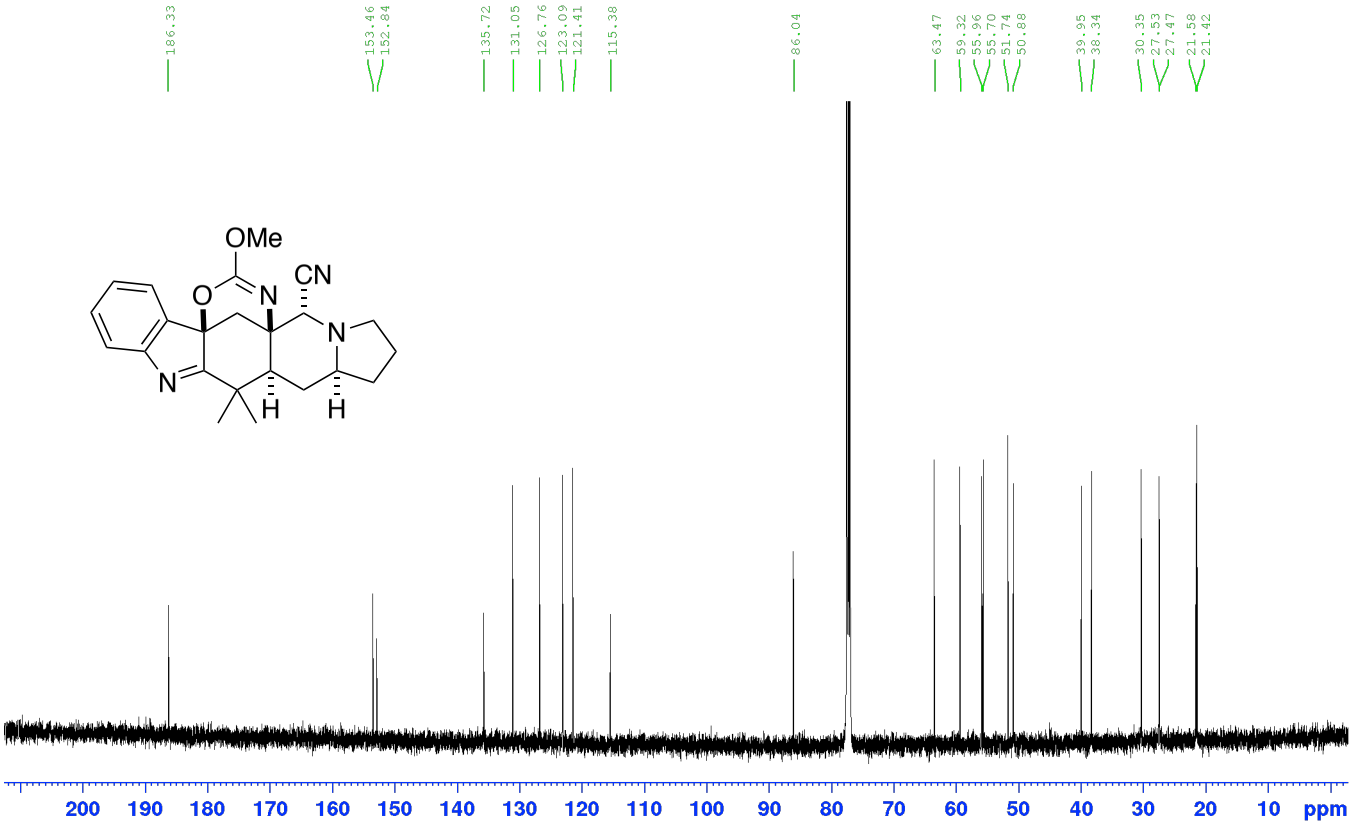
^{13}C NMR of compound **58** (125 MHz, $CDCl_3$)



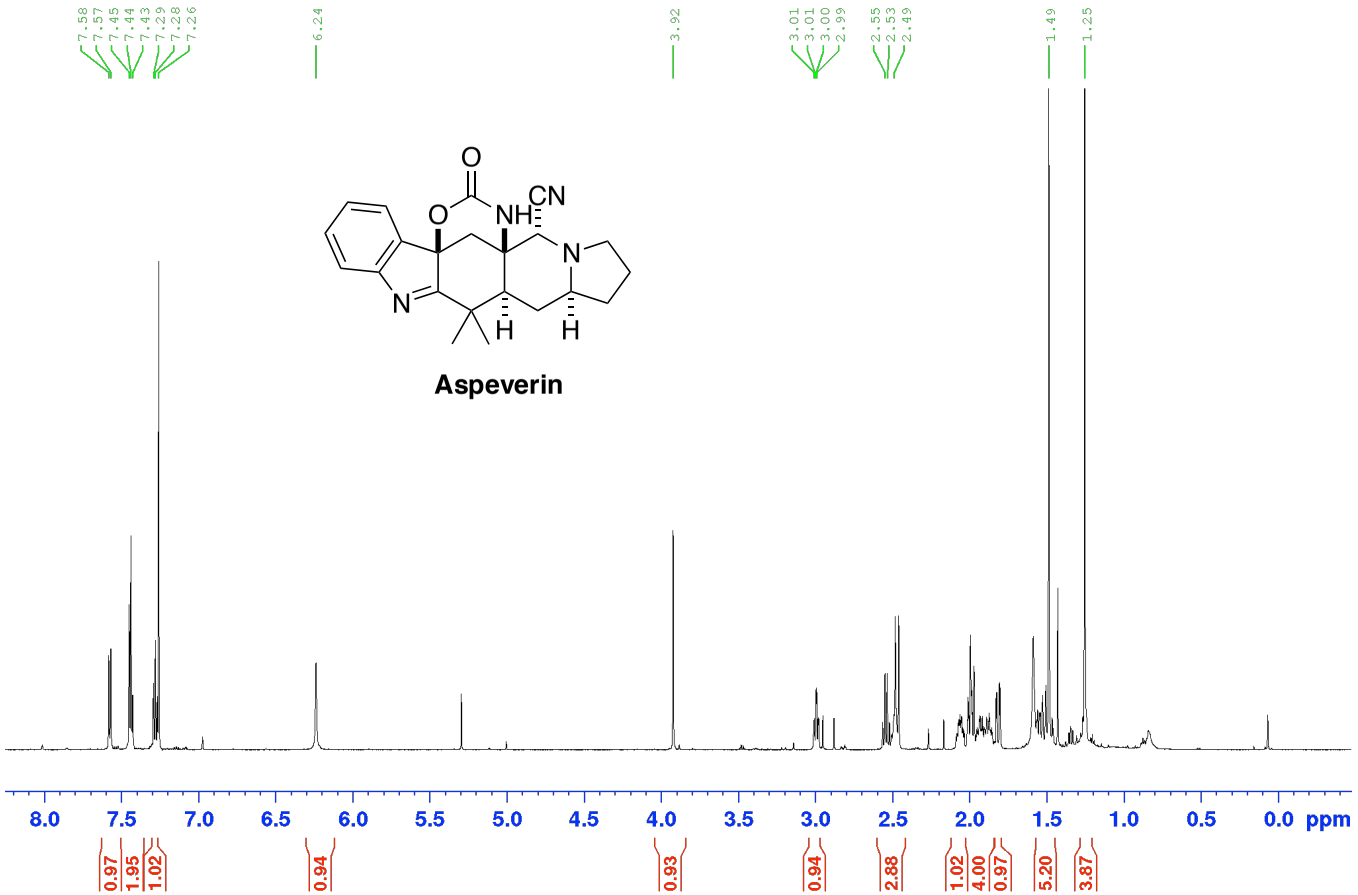
HSQC spectrum of compound **58** (500 MHz, $CDCl_3$)



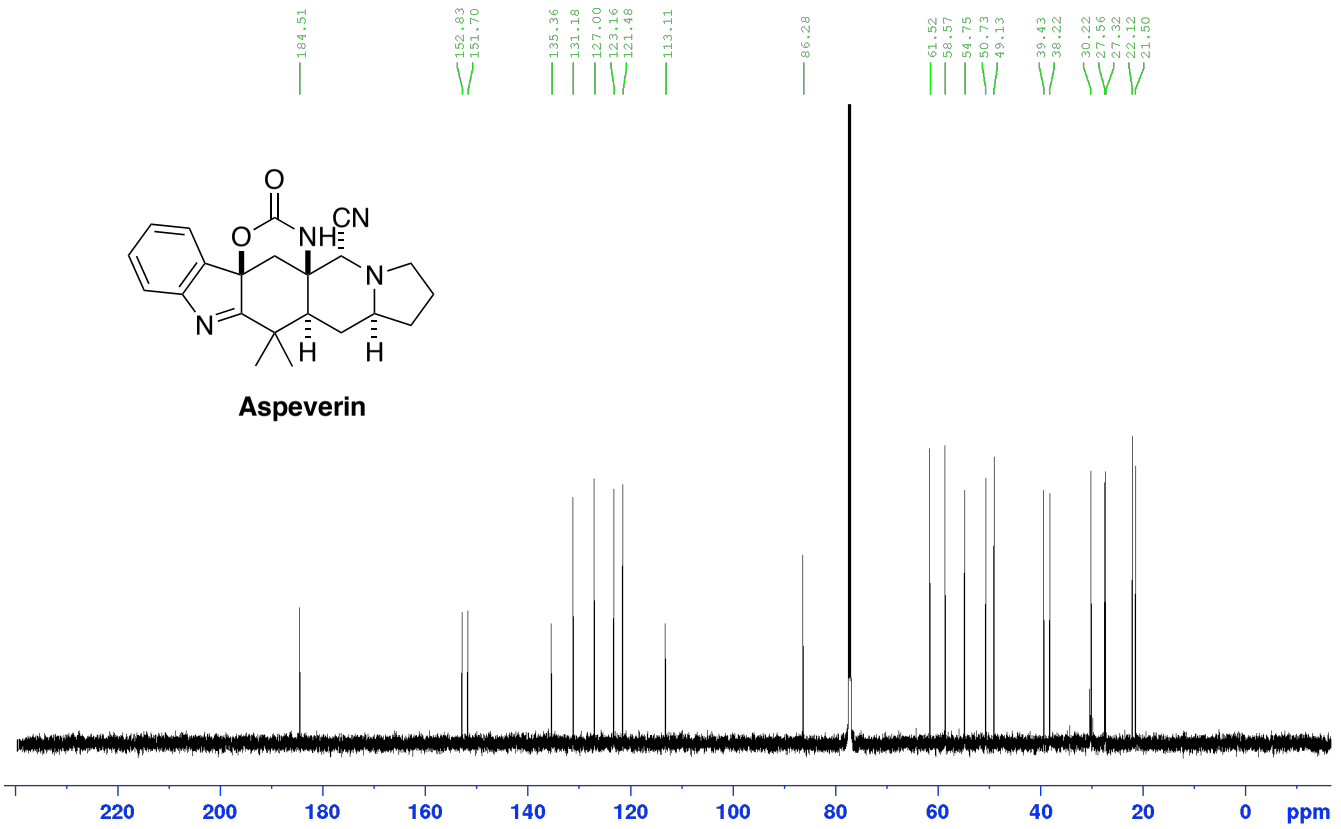
^{1}H NMR of compound **S-2** (500 MHz, $CDCl_3$)



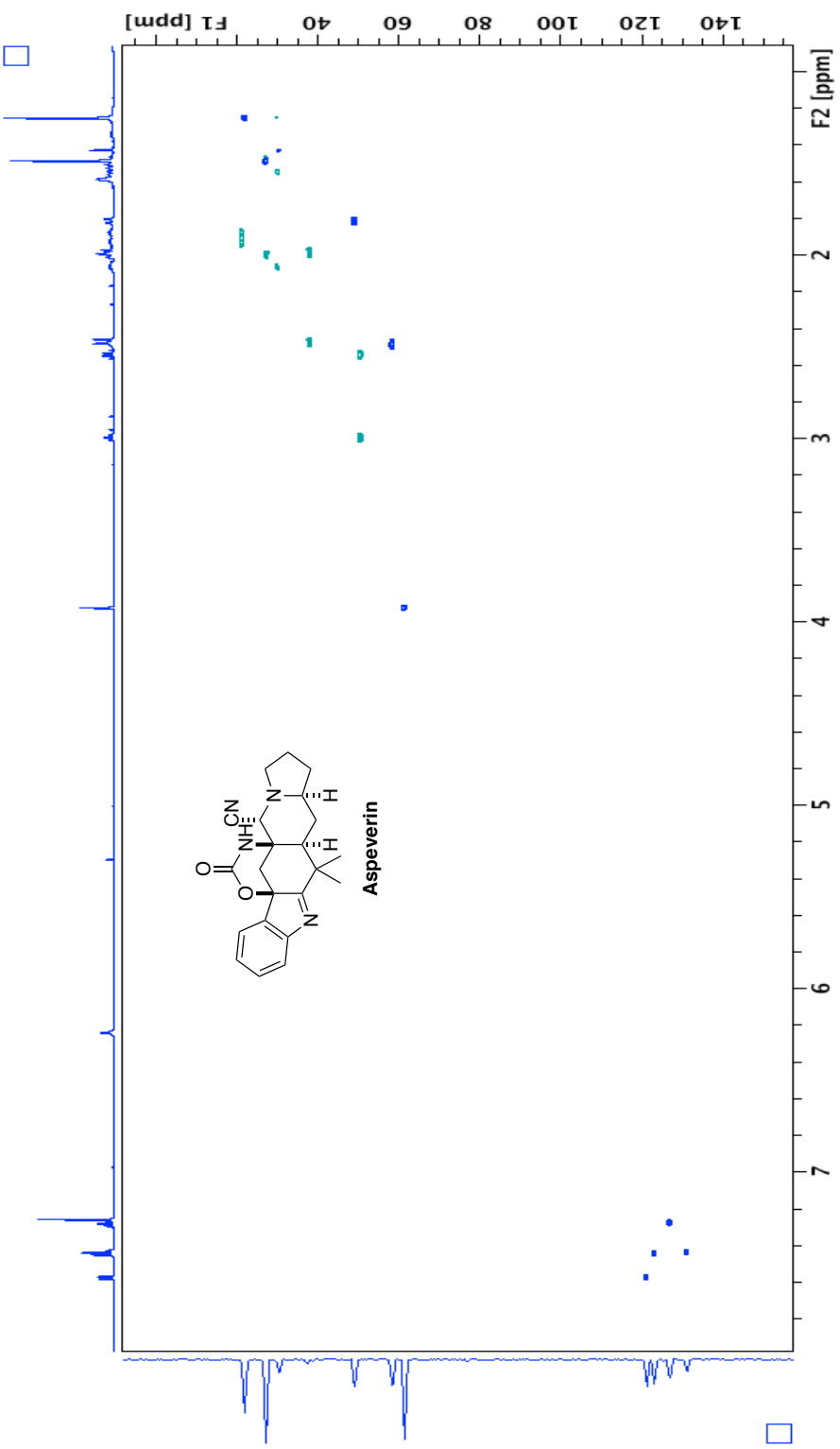
^{13}C NMR of compound **S-2** (125 MHz, $CDCl_3$)



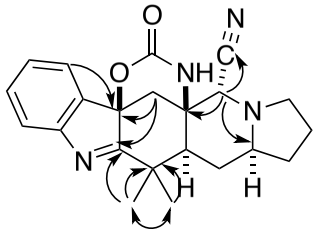
^{1}H NMR of compound **1** (600 MHz, $CDCl_3$)



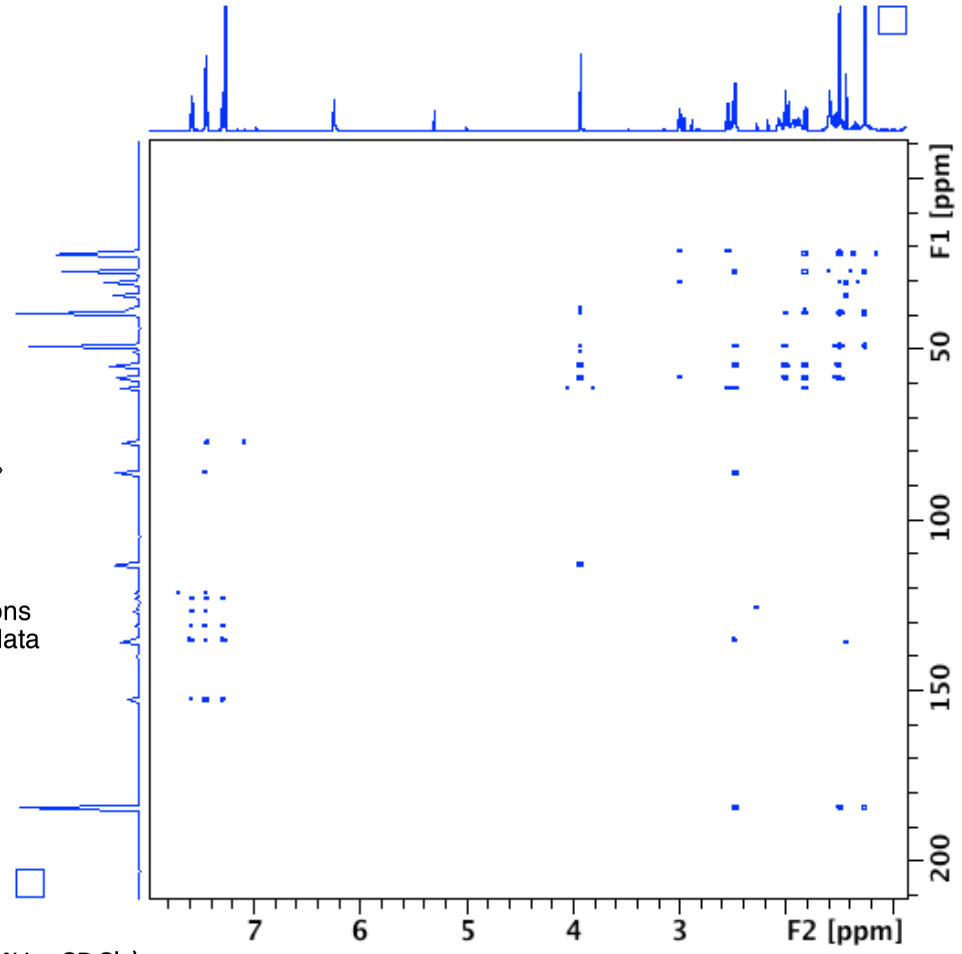
^{13}C NMR of compound **1** (150 MHz, $CDCl_3$)



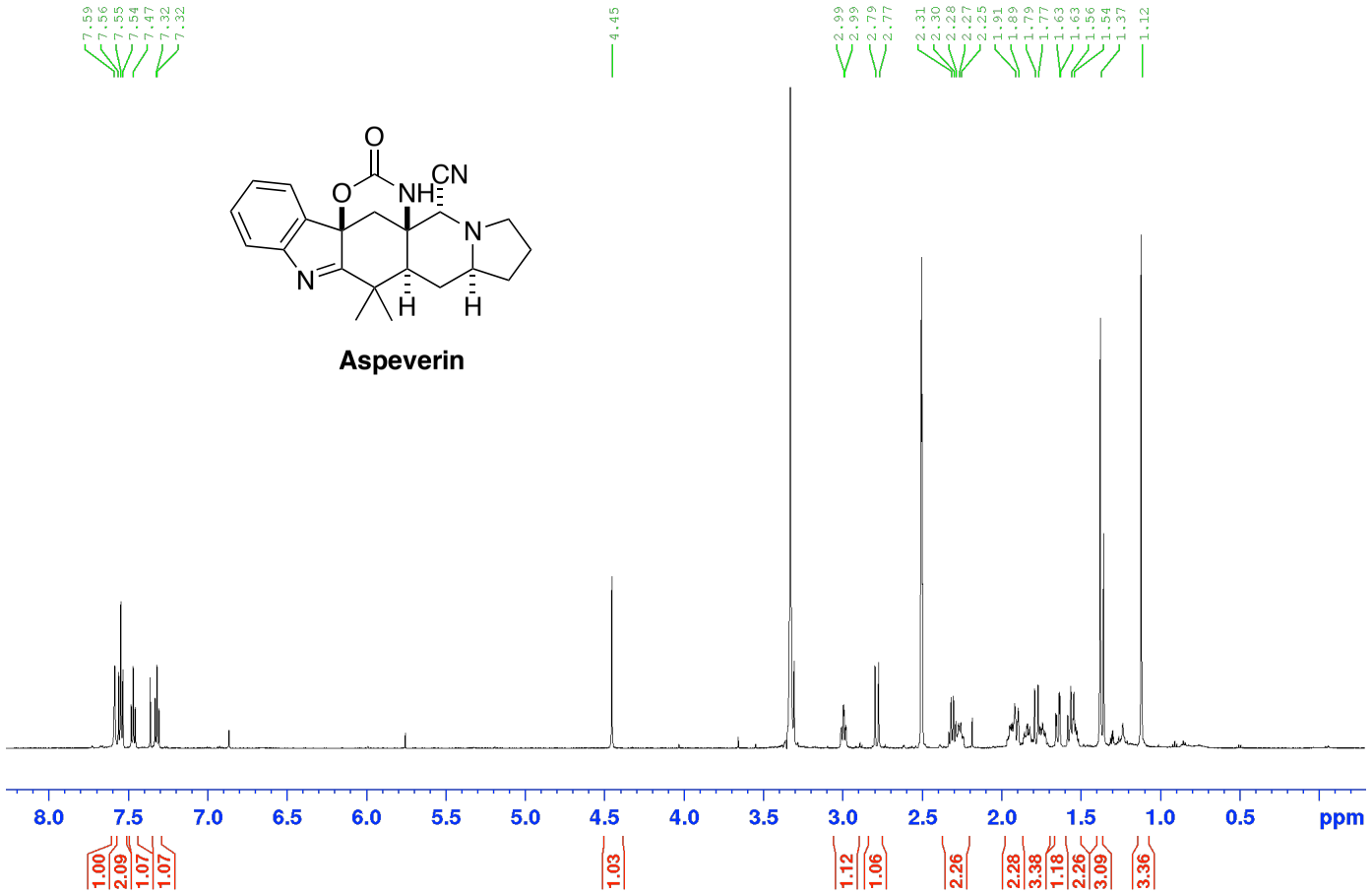
HSQC of compound **1** (600 MHz, $CDCl_3$)



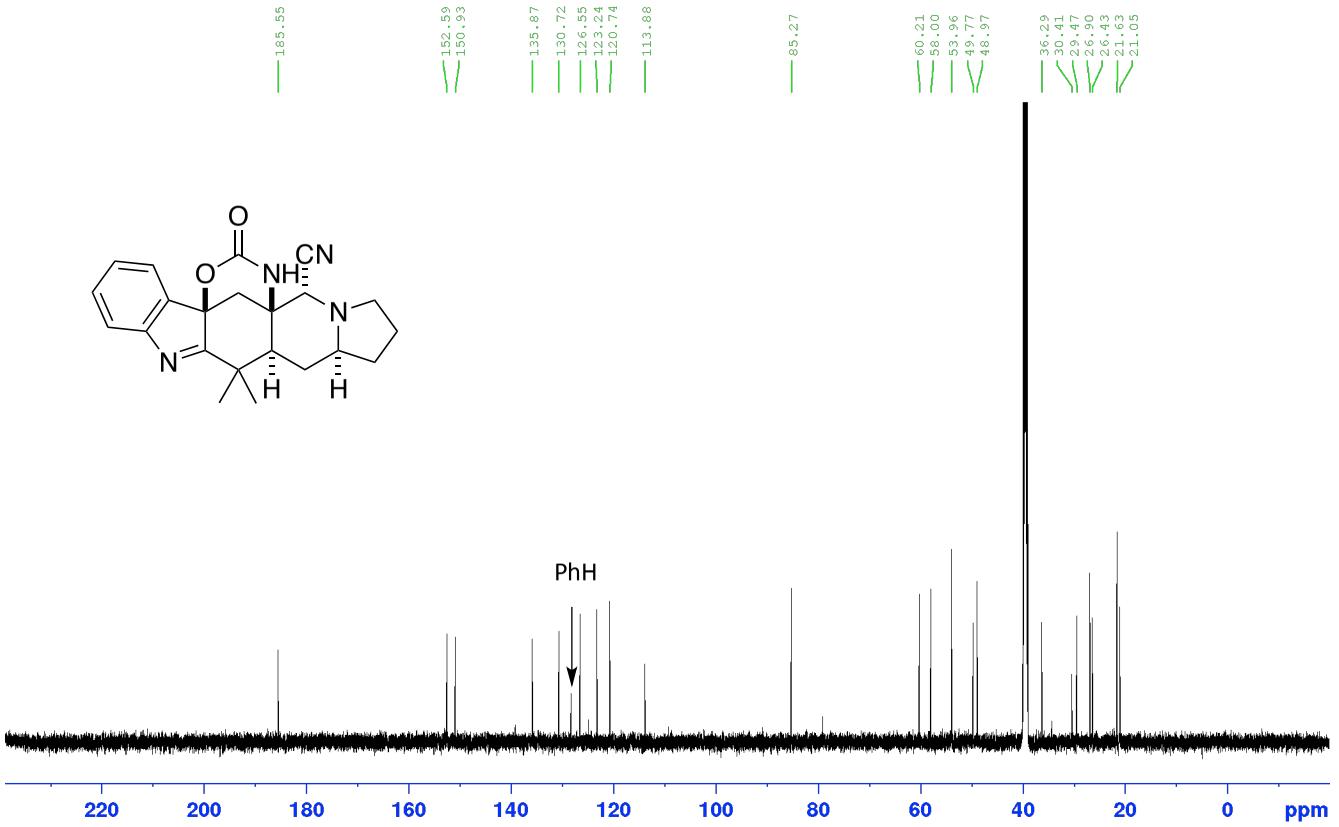
highlighted key correlations
consistent with reported data



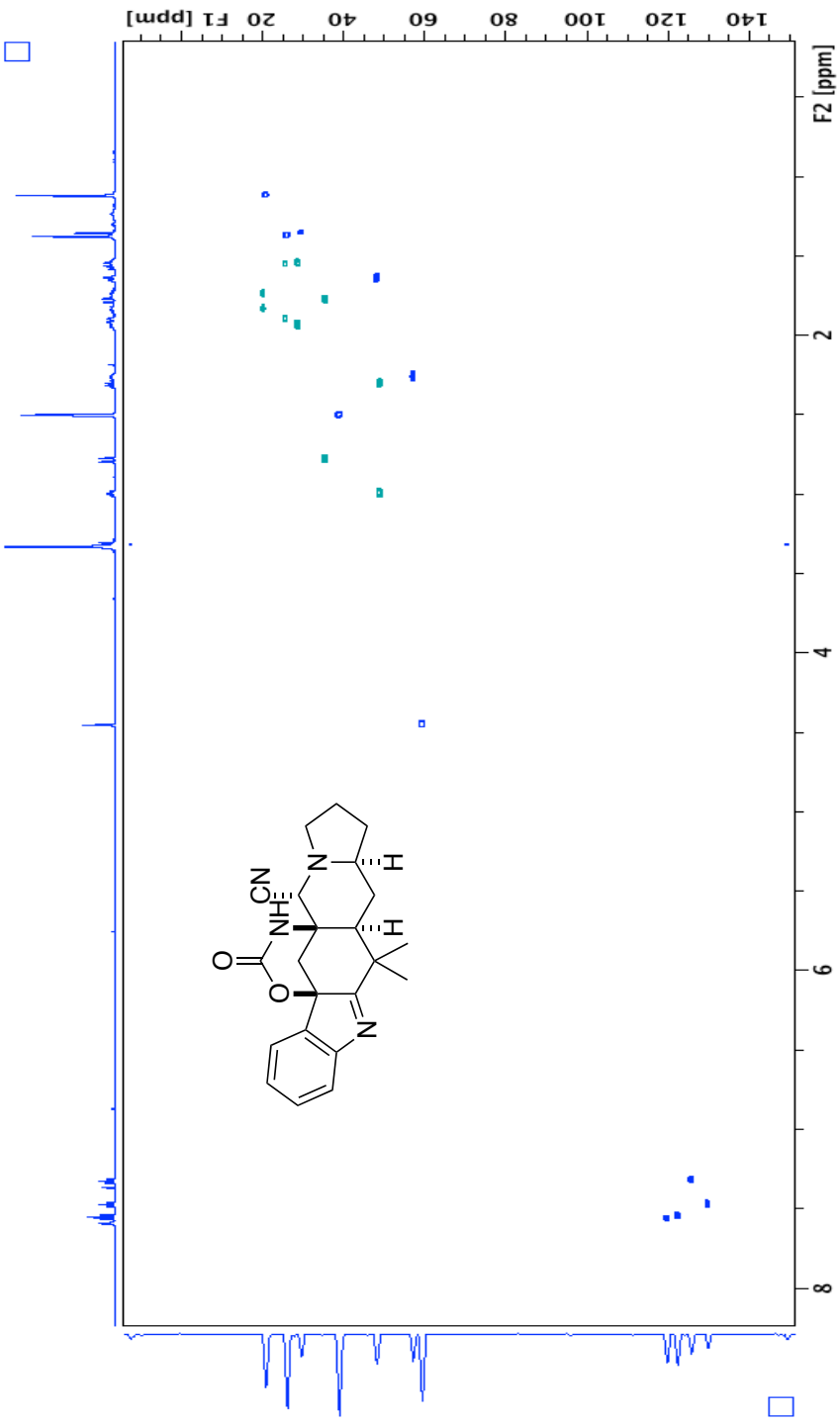
HMBC of compound **1** (600 MHz, $CDCl_3$)



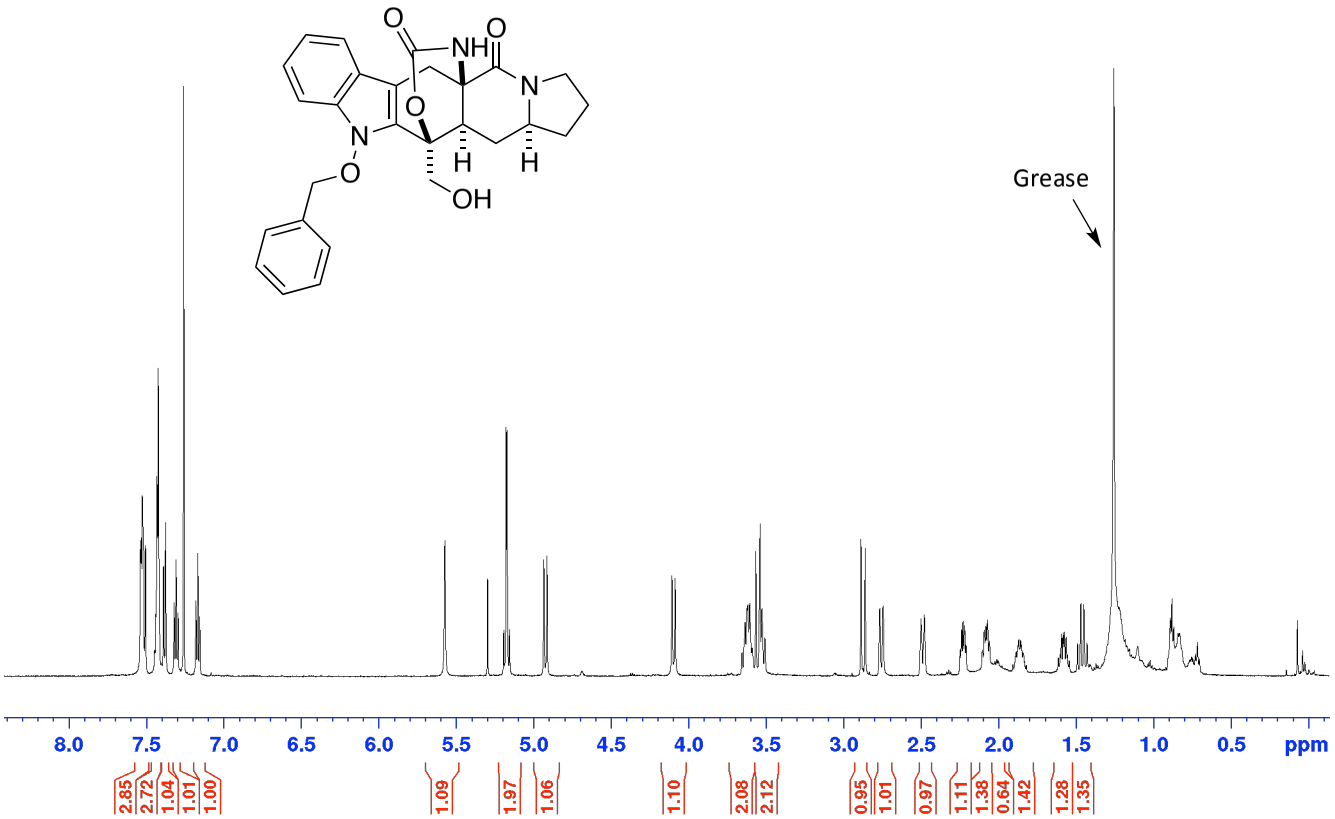
^{1}H NMR of compound **1** (600 MHz, DMSO-d_6)



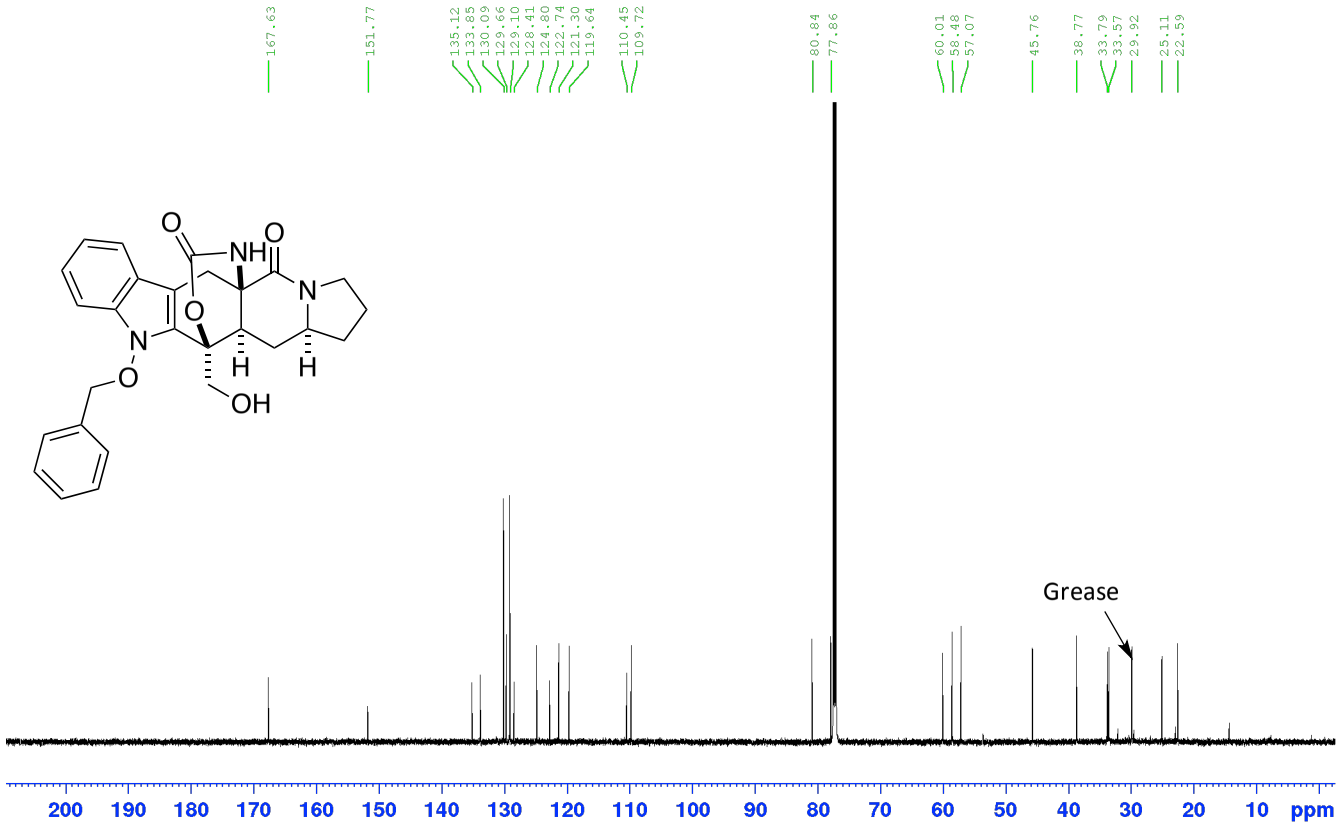
^{1}H NMR of compound **1** (600 MHz, DMSO-d_6)



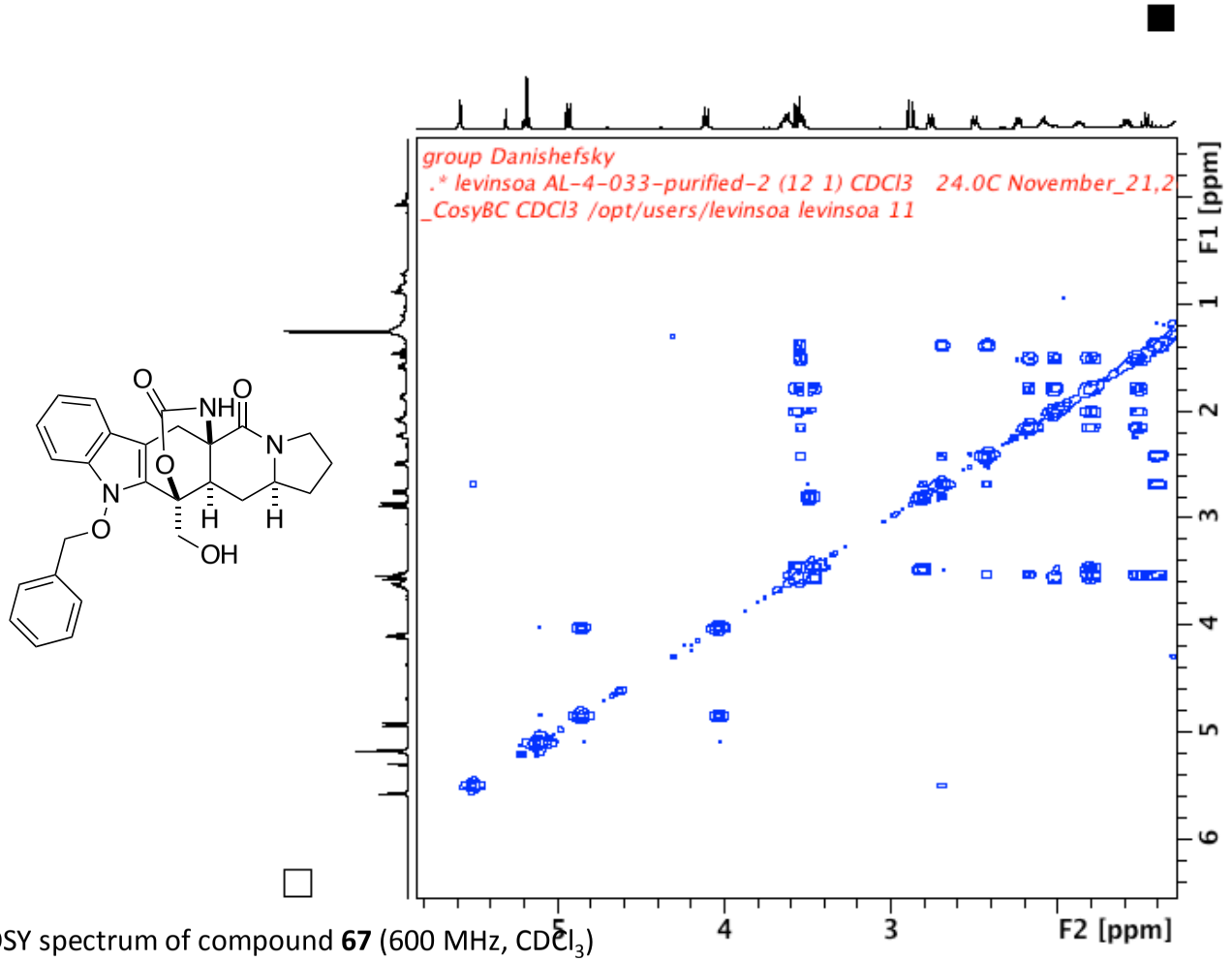
HSQC of compound **1** (600 MHz, DMSO-d_6)



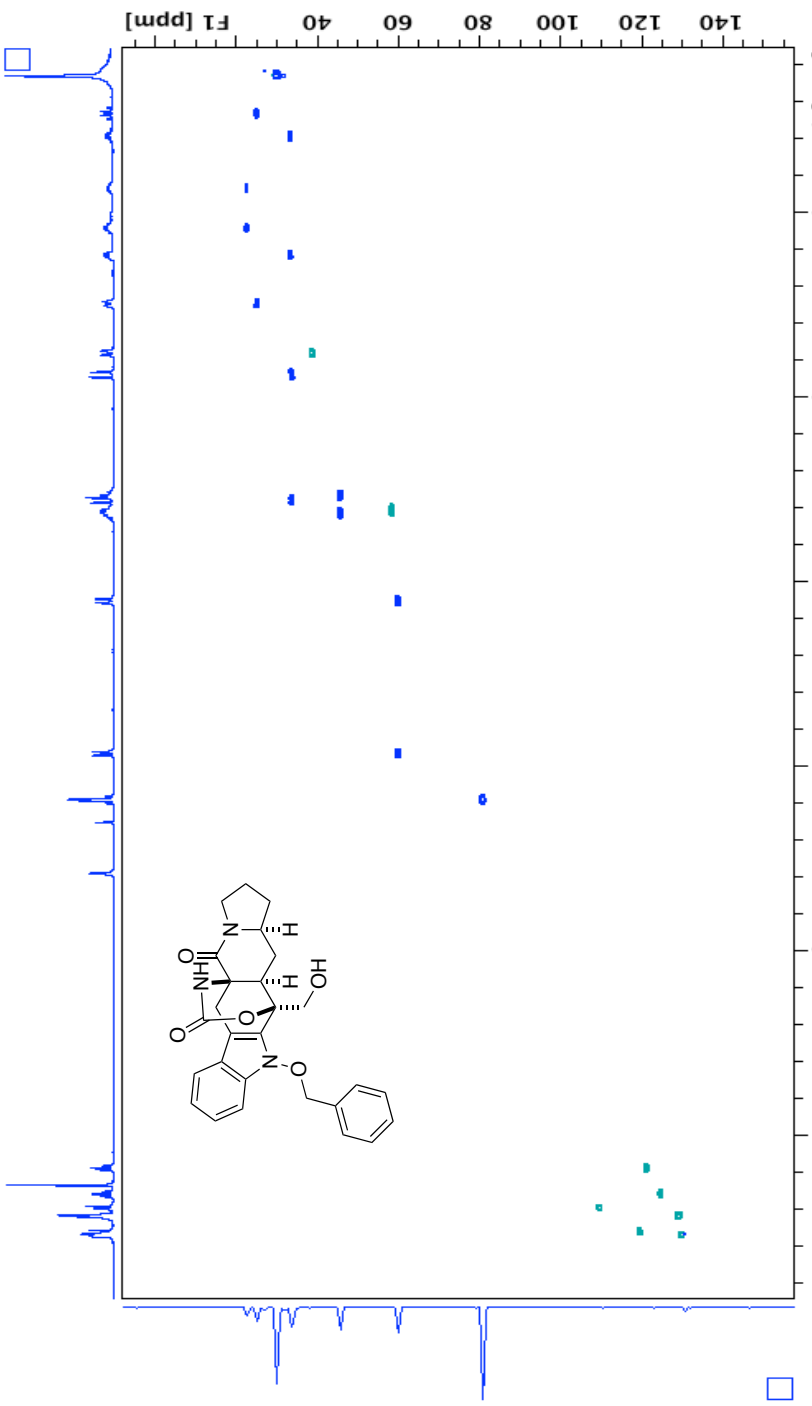
^{1}H NMR of compound **67** (600 MHz, $CDCl_3$)



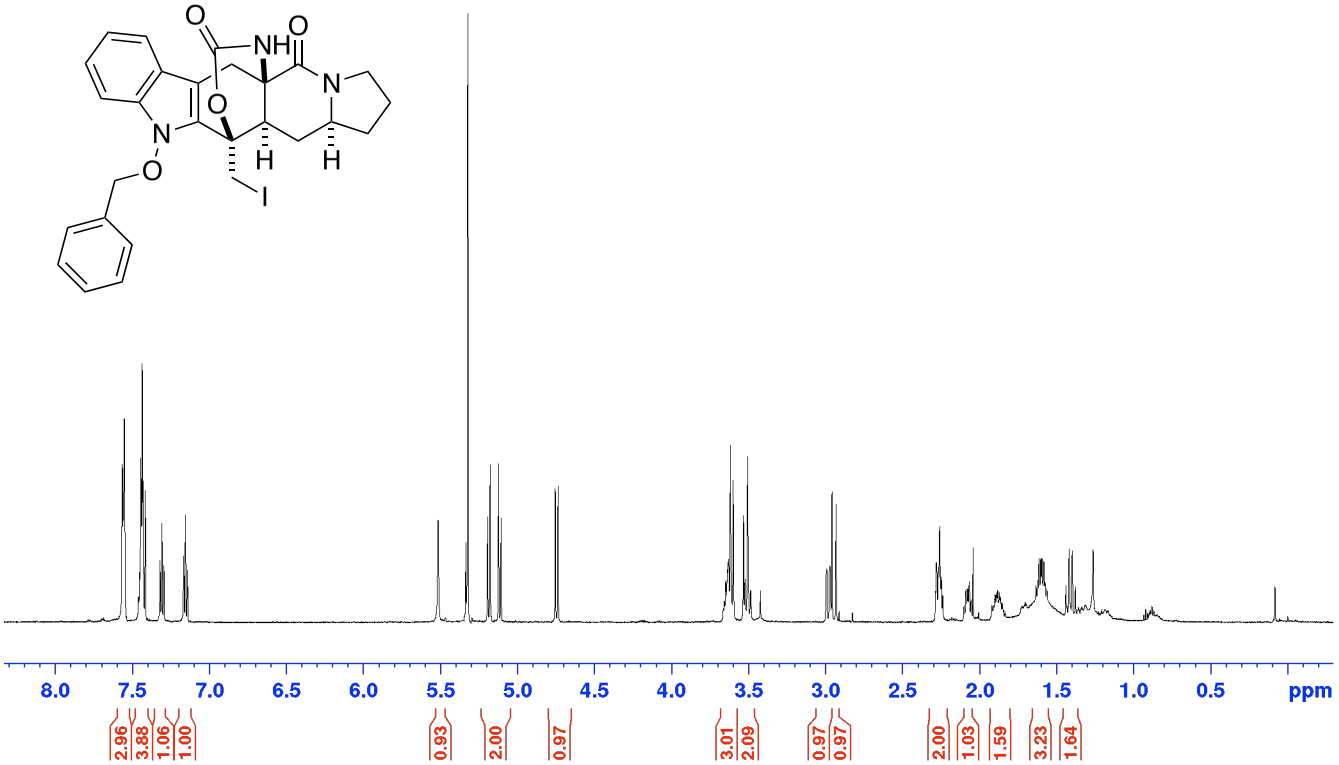
^{1}H NMR of compound **67** (600 MHz, $CDCl_3$)



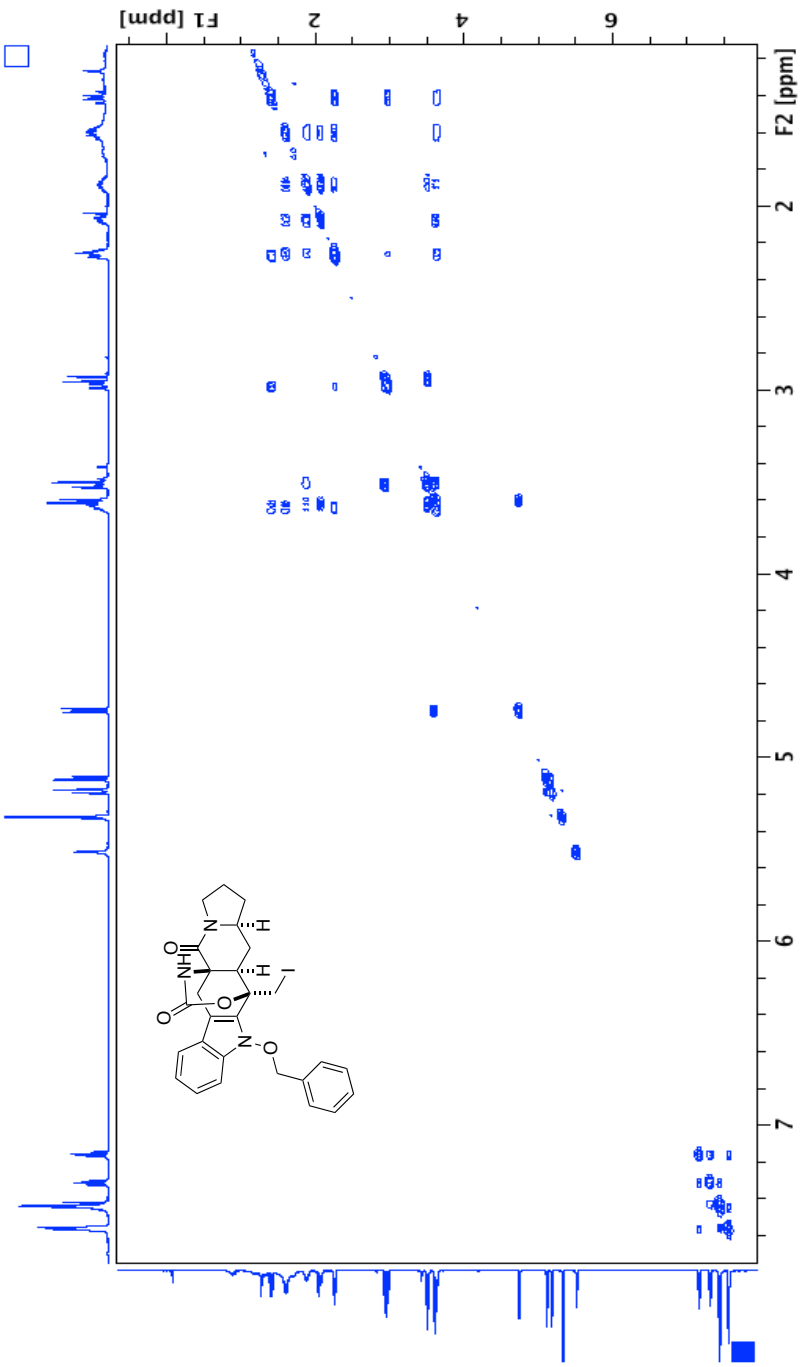
COSY spectrum of compound **67** (600 MHz, $CDCl_3$)



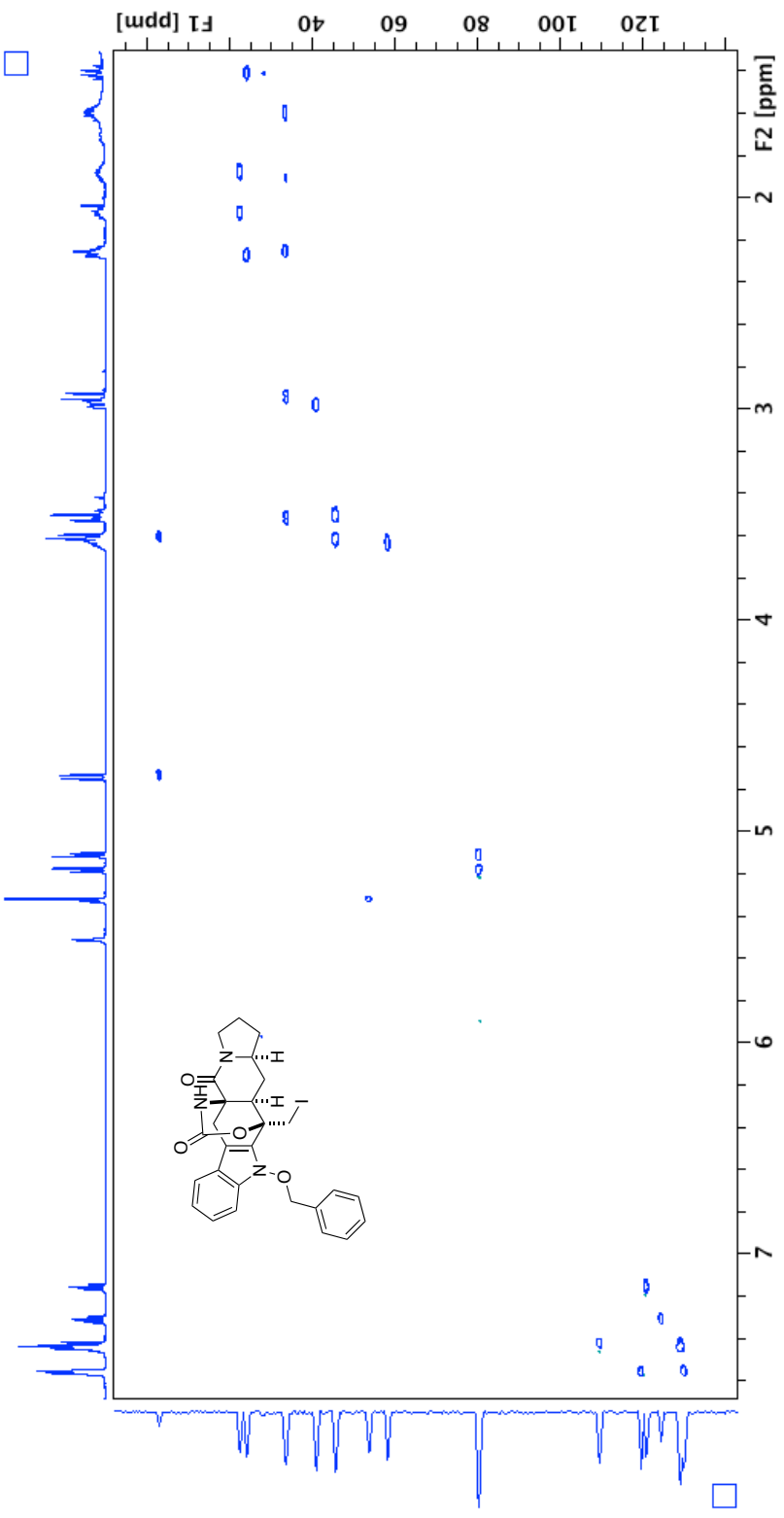
HSQC spectrum of compound **67** (600 MHz, $CDCl_3$)



^{1}H NMR of compound **68** (600 MHz, CD_2Cl_2)



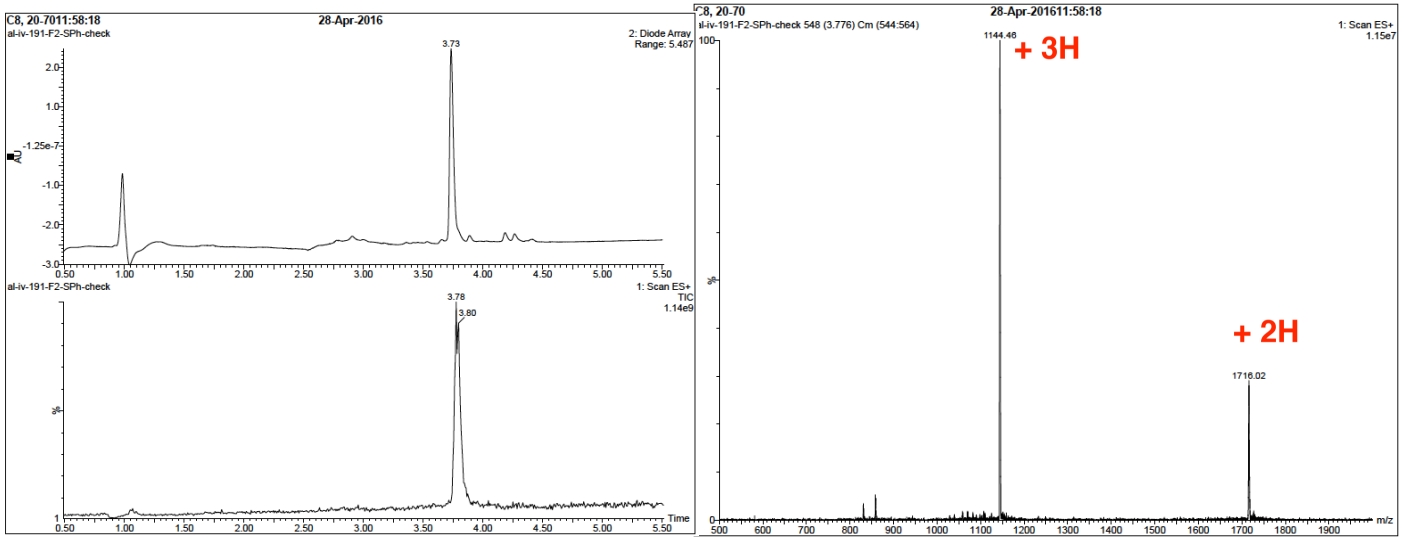
COSY spectrum of compound **68** (600 MHz, CD_2Cl_2)



HSQC spectrum of compound **68** (600 MHz, CD_2Cl_2)

Appendix II :
Selected UPLC/MS Spectra for Chapter 4

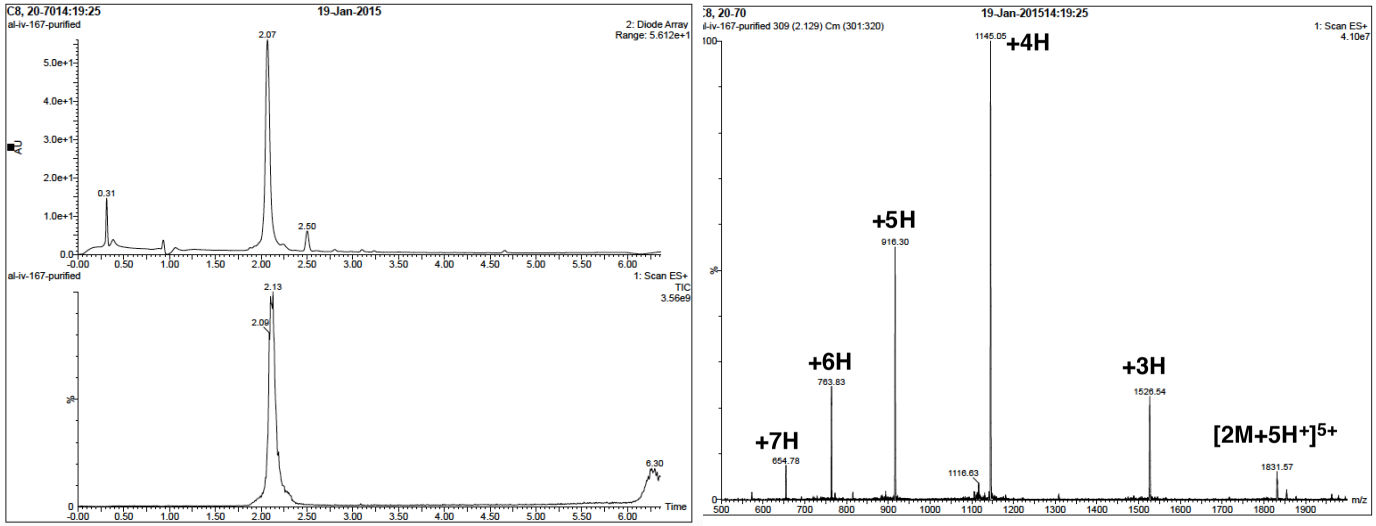
UPLC/MS and ESI(+)-MS Spectra of **2**



Chemical Formula: $C_{149}H_{225}N_{37}O_{48}S_4$
Molecular Weight: 3430.89

KRas [51-79] **(2)**

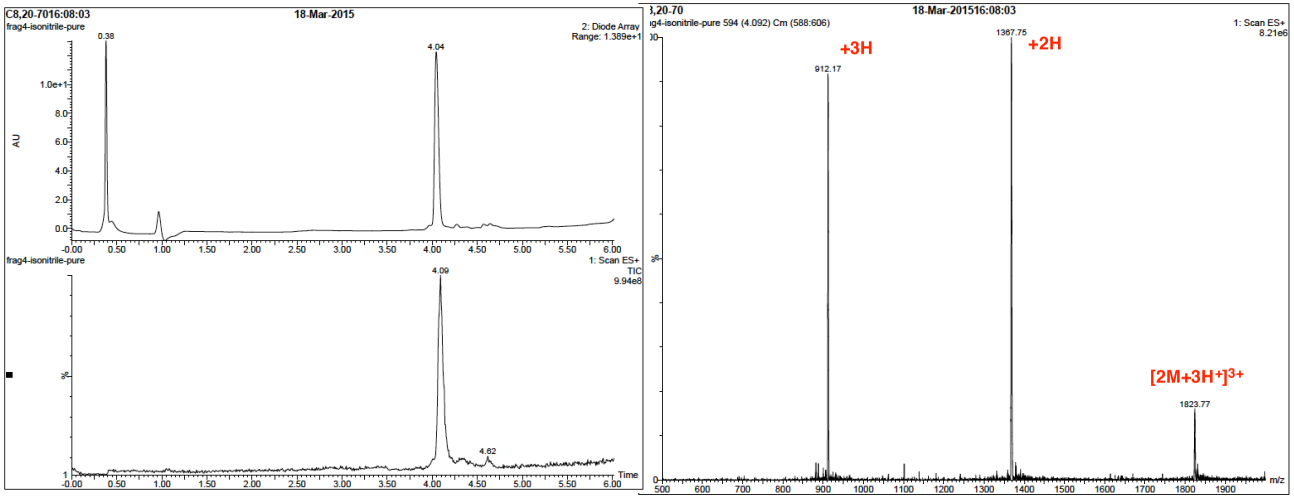
UPLC/MS and ESI(+)-MS Spectra of **3**



Chemical Formula: $C_{201}H_{320}N_{56}O_{60}S_3$
Molecular Weight: 4577.28

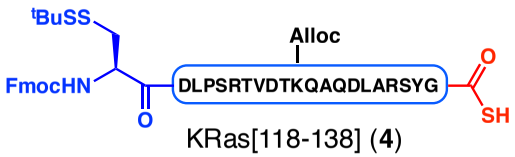
KRas [80-117] (**3**)

UPLC/MS and ESI(+)-MS Spectra of **4**



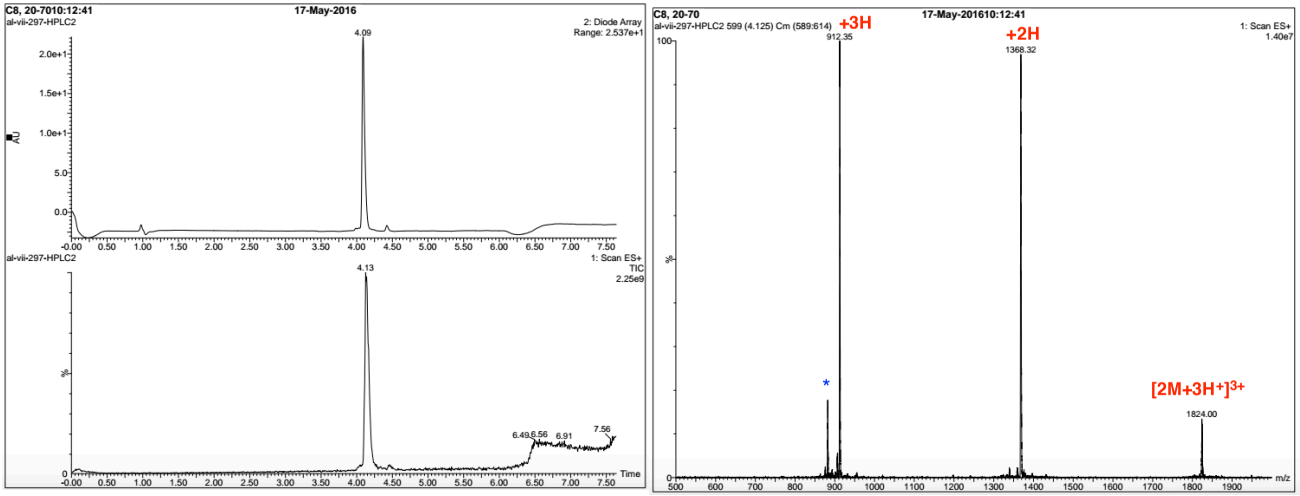
Chemical Formula: $C_{119}H_{180}N_{30}O_{38}S_3$
Molecular Weight: 2735.10

(all-L residues)



KRas[118-138] (**4**)

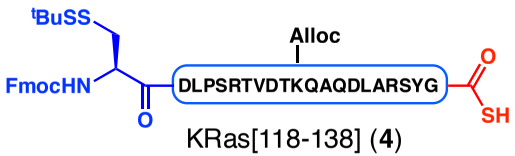
UPLC/MS and ESI(+)-MS Spectra of **4**



*Fragmentation mass of loss of -S^{t}Bu

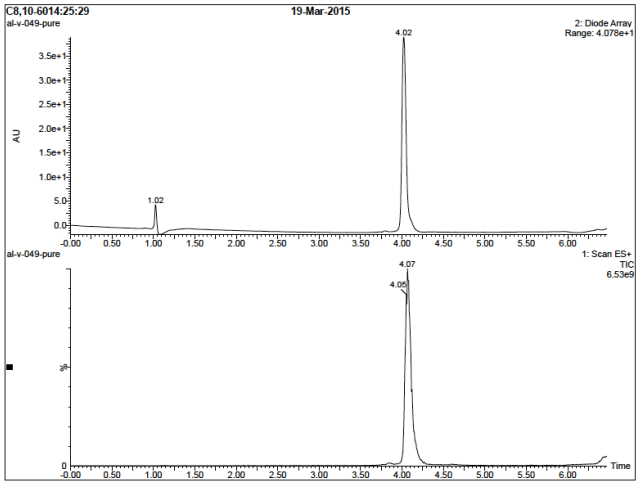
Chemical Formula: $C_{119}H_{180}N_{30}O_{38}S_3$
Molecular Weight: 2735.10

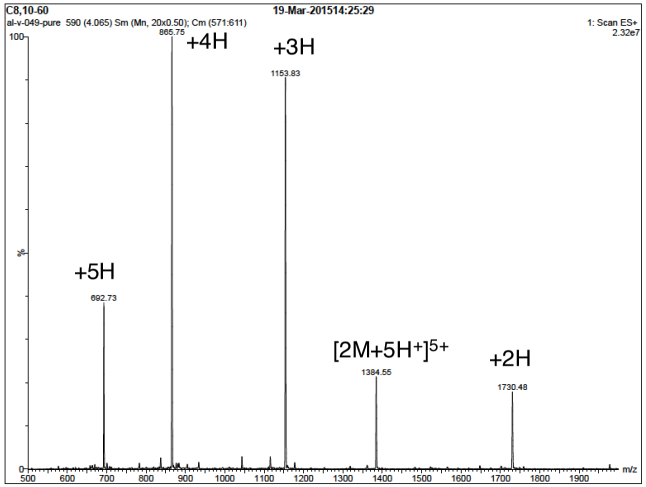
(all-D residues)



KRas[118-138] (**4**)

UPLC/MS and ESI(+)-MS Spectra of **5**

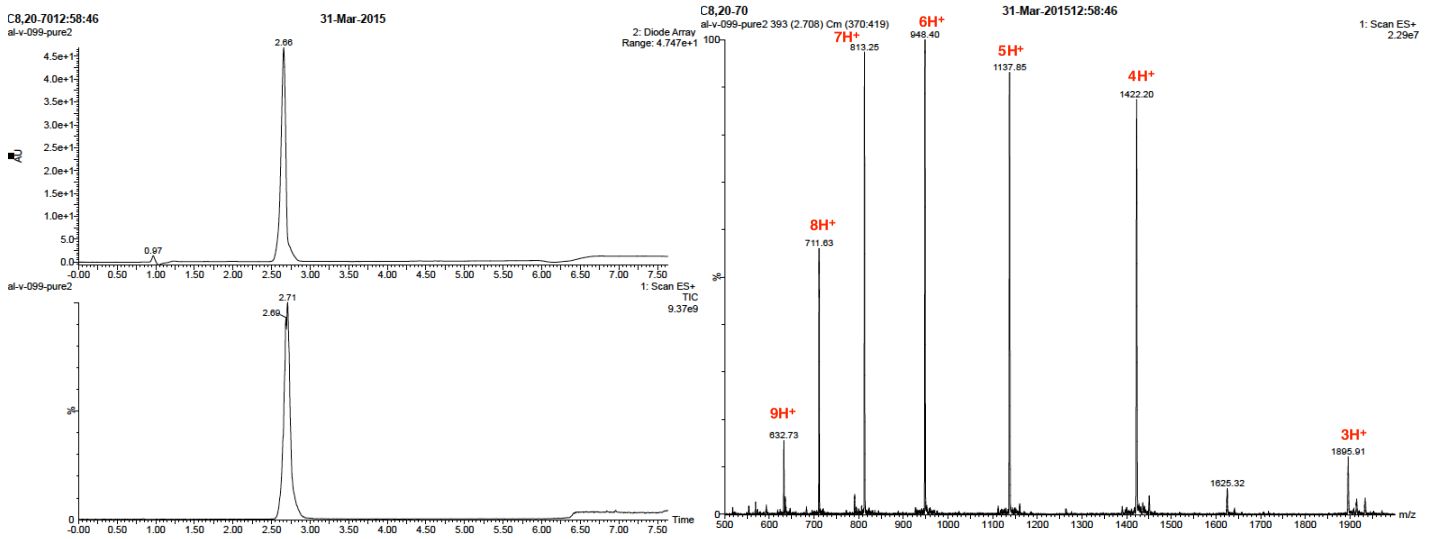




Chemical Formula: $C_{156}H_{244}N_{42}O_{47}$
Molecular Weight: 3459.92

KRas [139-166] **(5)**

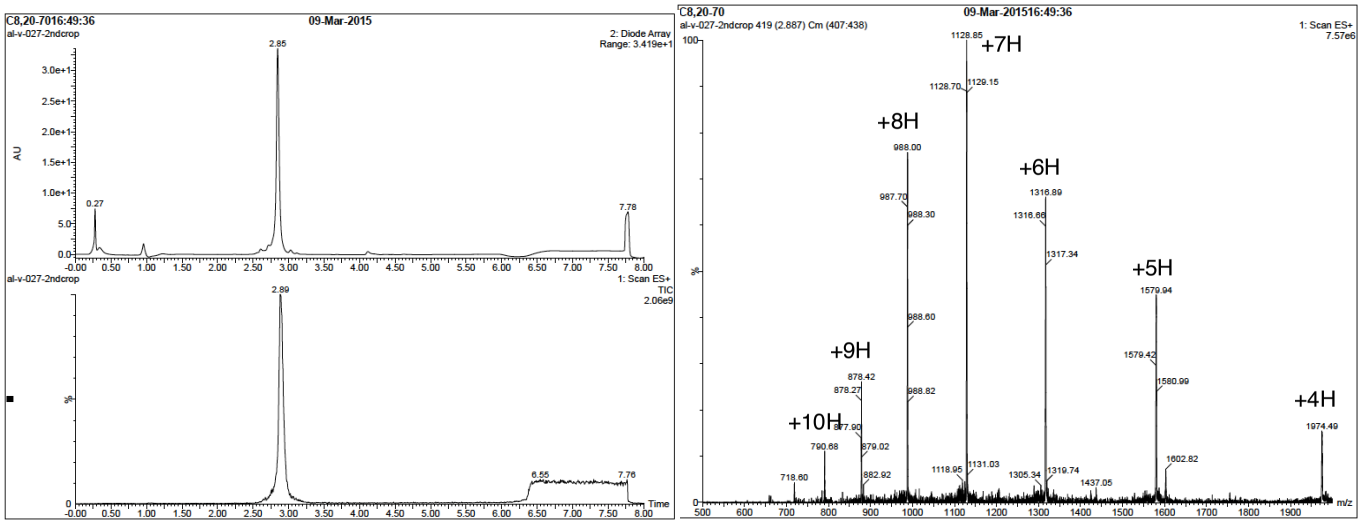
UPLC/MS and ESI(+)-MS Spectra of **10**



Chemical Formula: $C_{248}H_{400}N_{72}O_{77}S_2$
Molecular Weight: 5686.48

KRas [118-166] **(10)**

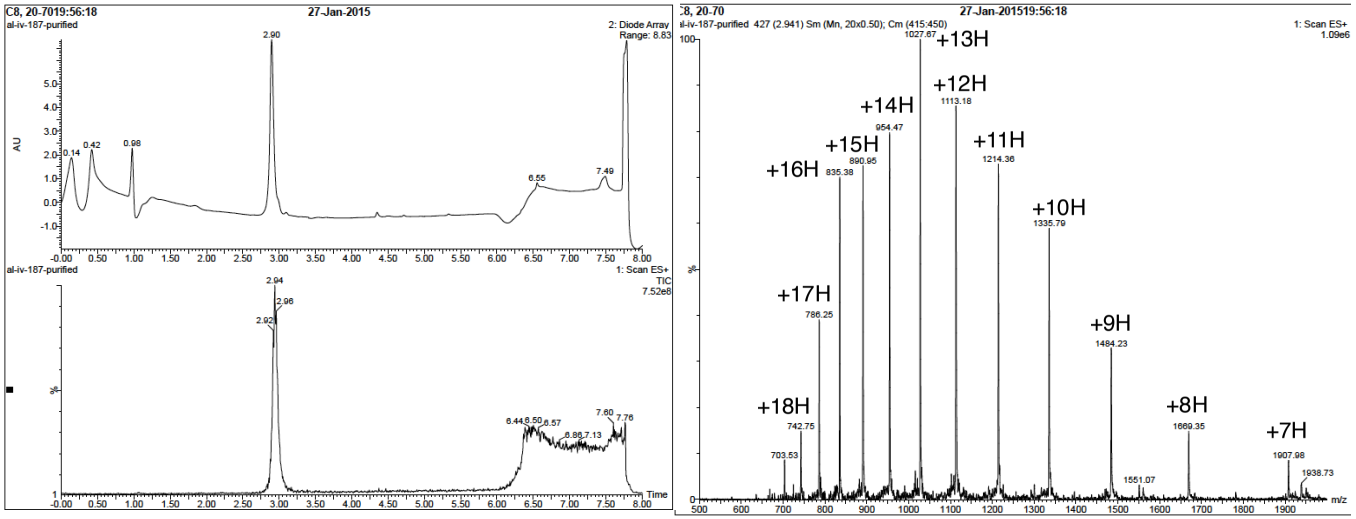
UPLC/MS and ESI(+)-MS Spectra of **11**



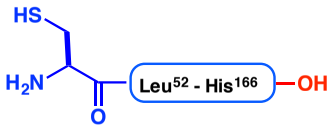
KRas [51-117] **(11)**

Chemical Formula: $C_{344}H_{539}N_{93}O_{108}S_6$
Molecular Weight: 7898.00

UPLC/MS and ESI(+)-MS Spectra of **12**

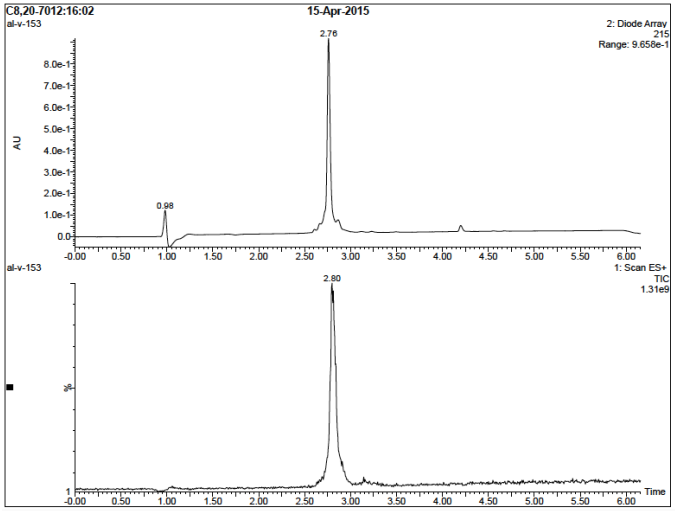


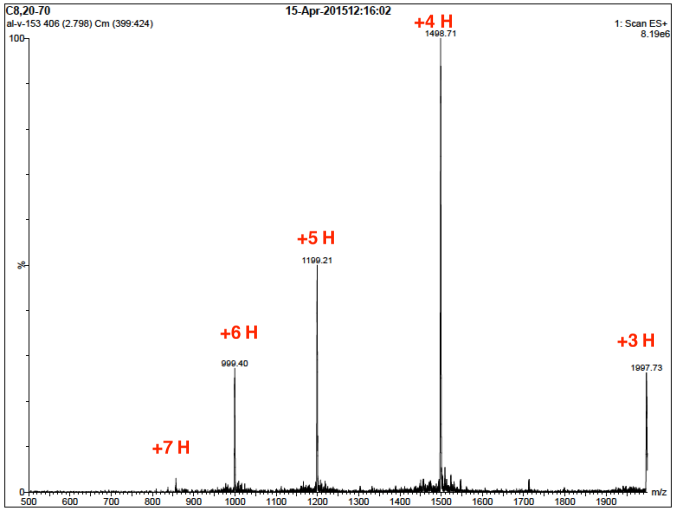
Chemical Formula: $C_{582}H_{921}N_{165}O_{183}S_6$
Molecular Weight: 13350.10



KRas [51-166] **(12)**

UPLC/MS and ESI(+)-MS Spectra of **14**



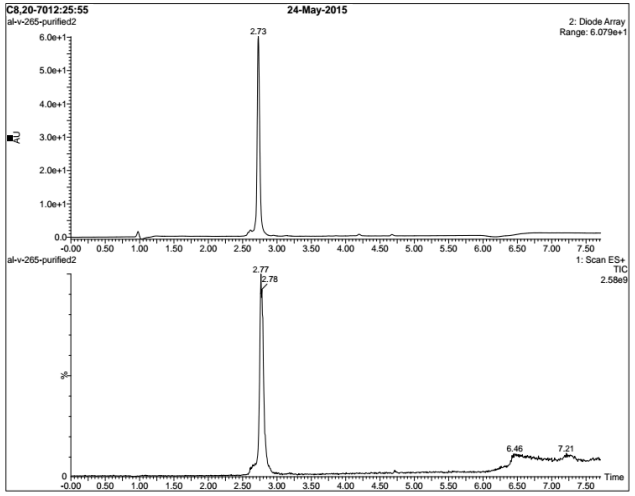


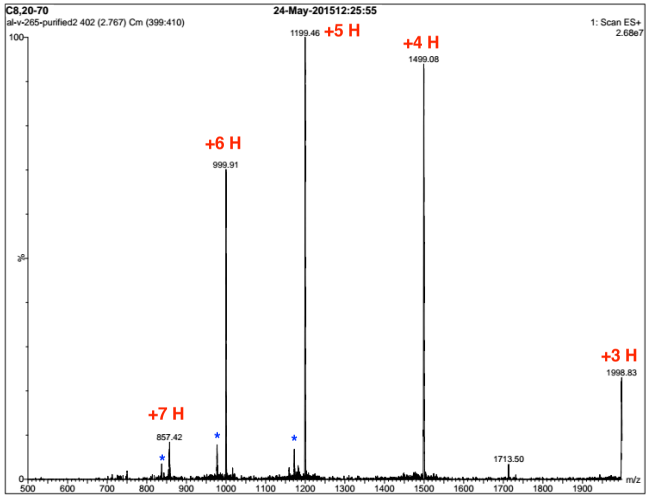
Chemical Formula: $C_{273}H_{418}N_{62}O_{85}S_2$
Molecular Weight: 5992.82

(all-L residues)

KRas[1-50] (**14**)

UPLC/MS and ESI(+)-MS Spectra of **14**





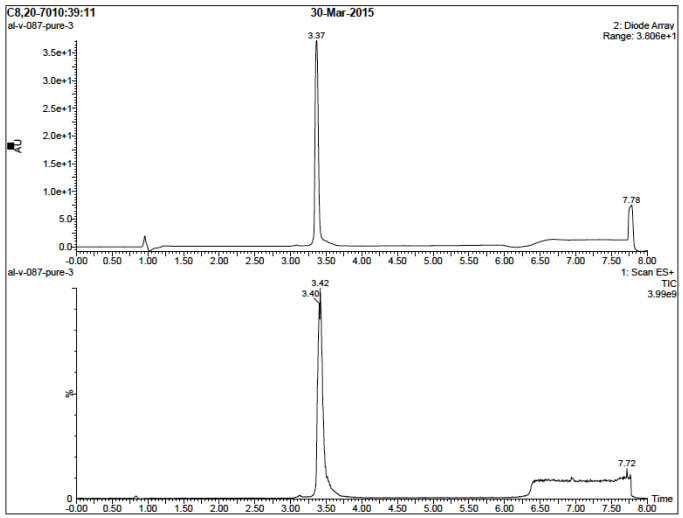
*denotes fragmentation mass of Hmb group

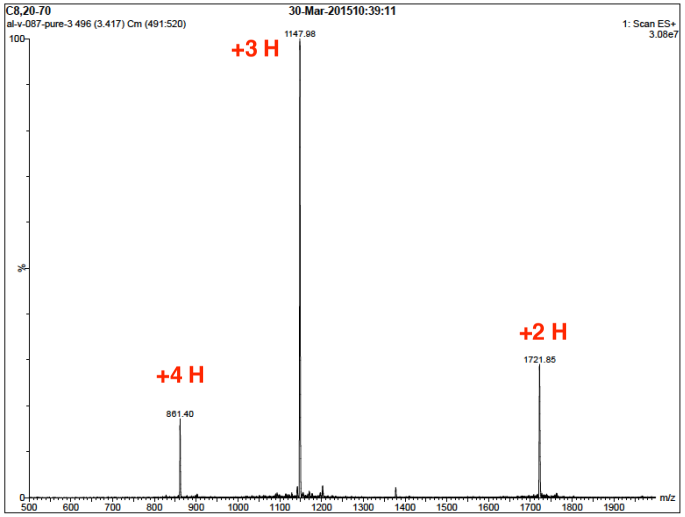
Chemical Formula: $C_{273}H_{418}N_{62}O_{85}S_2$
Molecular Weight: 5992.82

(all-D residues)

KRas[1-50] (**14**)

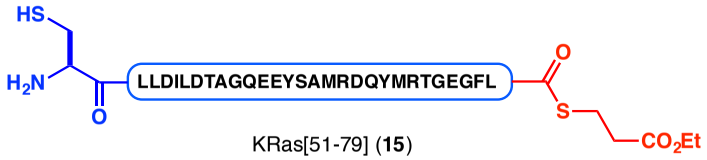
UPLC/MS and ESI(+)-MS Spectra of **15**





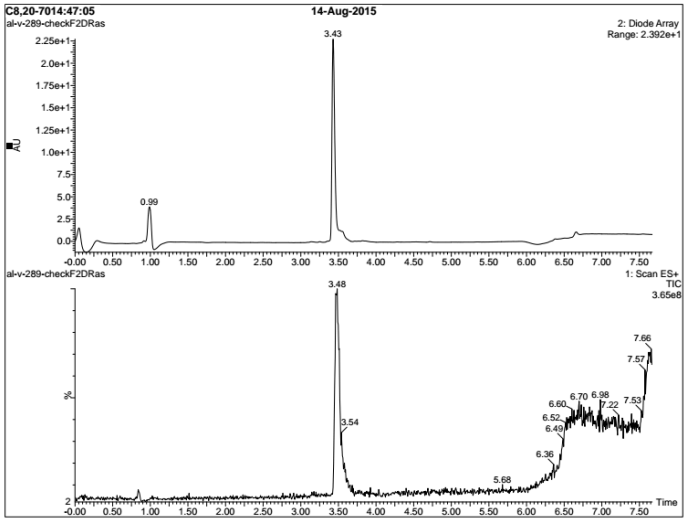
Chemical Formula: $C_{147}H_{229}N_{37}O_{50}S_4$
Molecular Weight: 3442.90

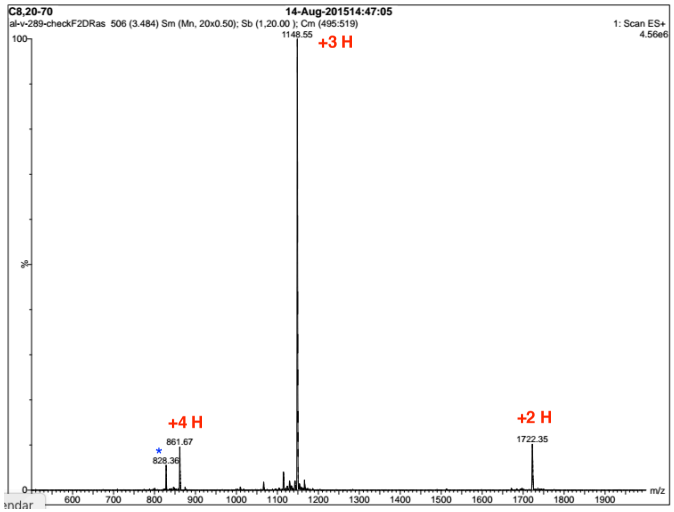
All-L Residues



KRas[51-79] (**15**)

UPLC/MS and ESI(+)-MS Spectra of **15**

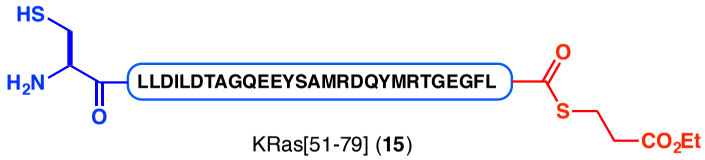




*Fragmentation mass of loss of thioester

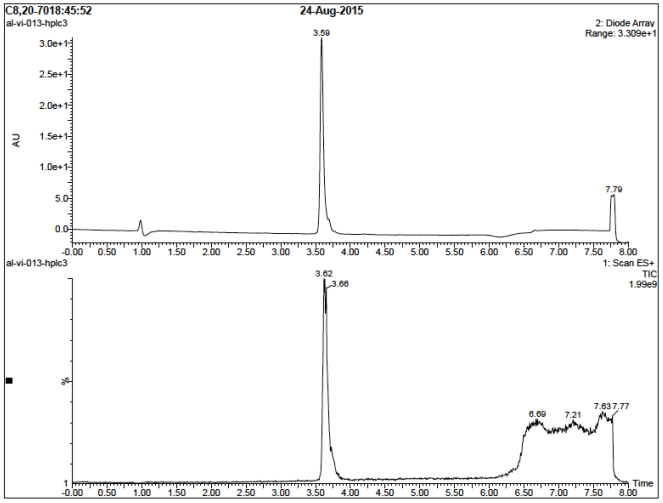
Chemical Formula: $C_{147}H_{229}N_{37}O_{50}S_4$
Molecular Weight: 3442.90

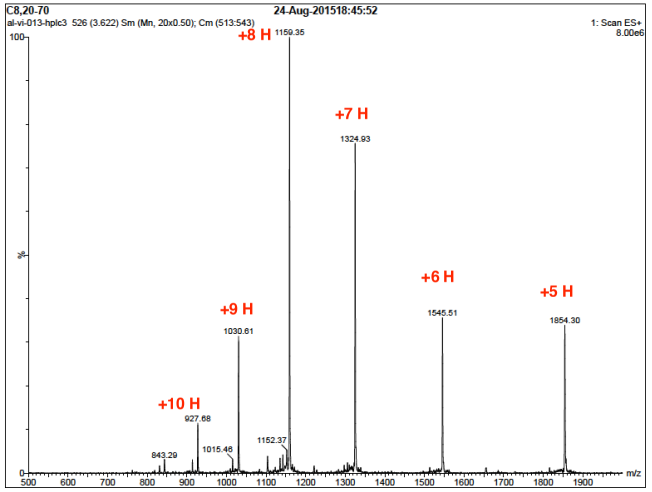
All-D Residues



KRas[51-79] (**15**)

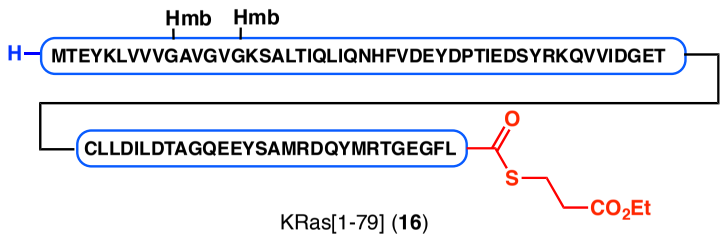
UPLC/MS and ESI(+)-MS Spectra of **16**





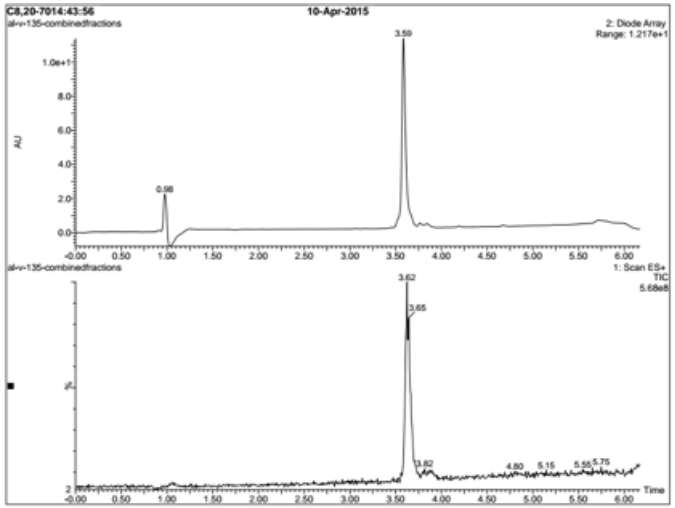
Chemical Formula: $C_{412}H_{639}N_{99}O_{133}S_5$
Molecular Weight: 9267.50

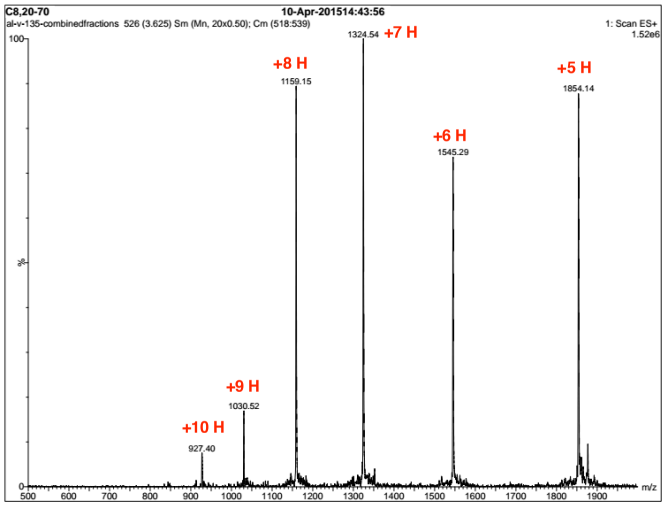
(All-L Residues)



KRas[1-79] (**16**)

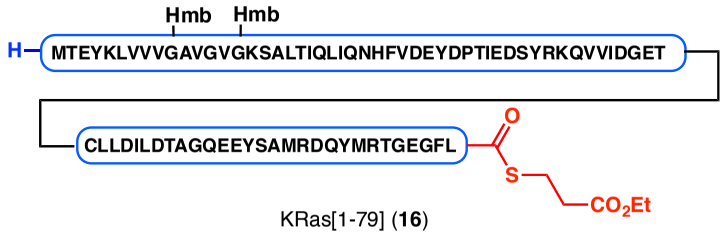
UPLC/MS and ESI(+)-MS Spectra of **16**





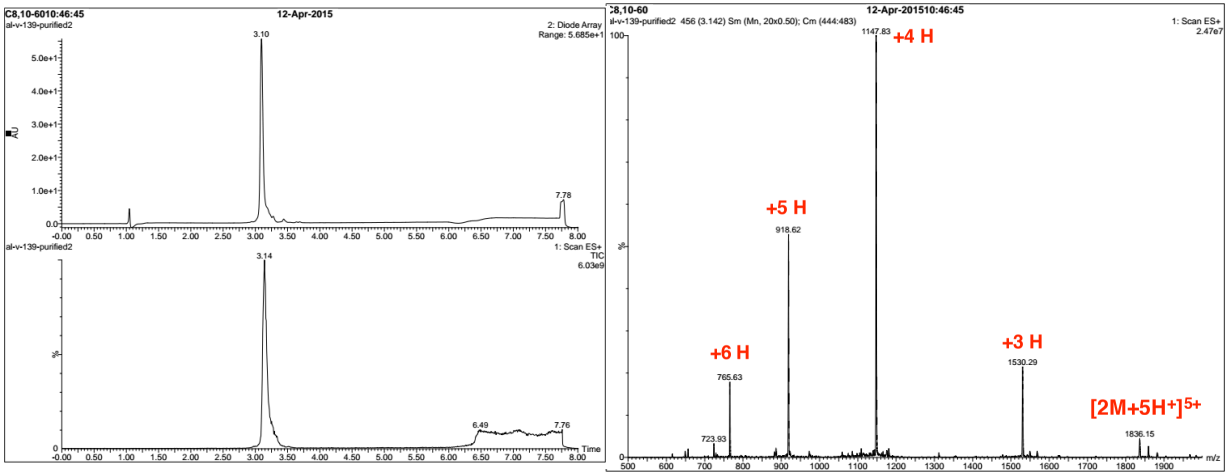
Chemical Formula: $C_{412}H_{639}N_{99}O_{133}S_5$
Molecular Weight: 9267.50

(all-D residues)



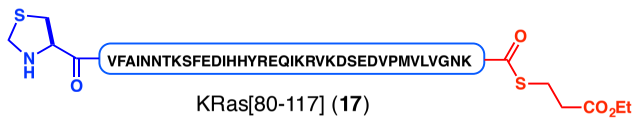
KRas[1-79] (**16**)

UPLC/MS and ESI(+)-MS Spectra of **17**



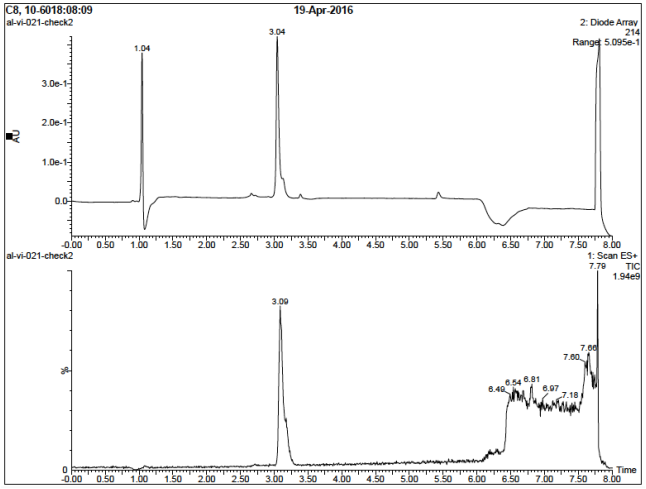
Chemical Formula: $C_{202}H_{320}N_{56}O_{60}S_3$
Molecular Weight: 4589.29

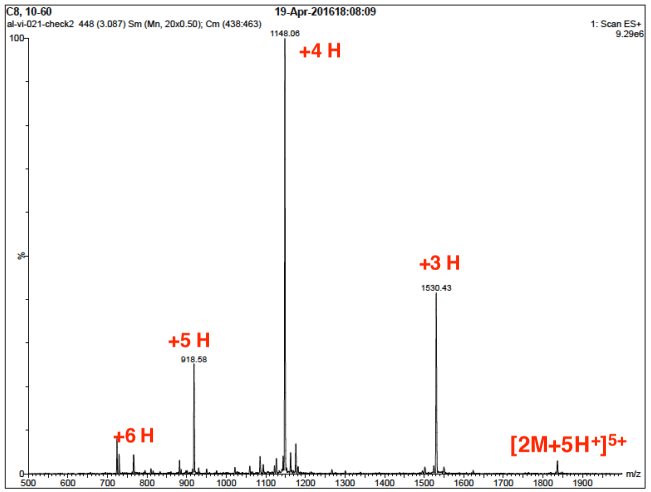
(All L-Residues)



KRas[80-117] (**17**)

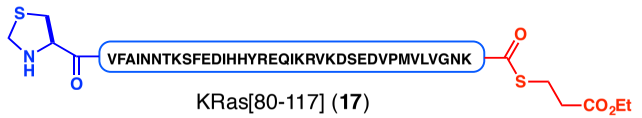
UPLC/MS and ESI(+)-MS Spectra of **17**





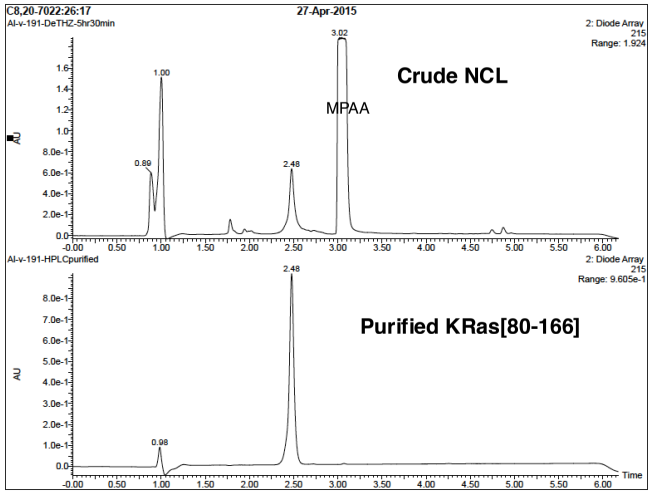
Chemical Formula: $C_{202}H_{320}N_{56}O_{60}S_3$
Molecular Weight: 4589.29

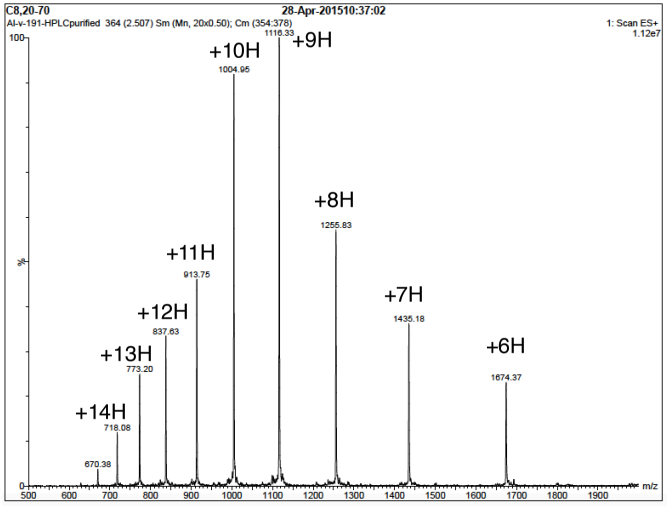
(all-D residues)



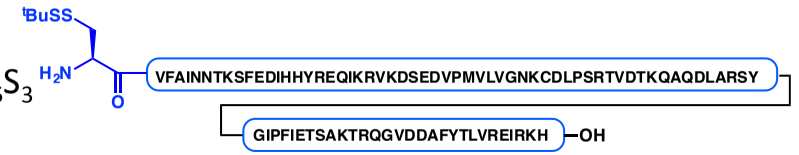
KRas[80-117] (**17**)

UPLC/MS and ESI(+)-MS Spectra of **18**



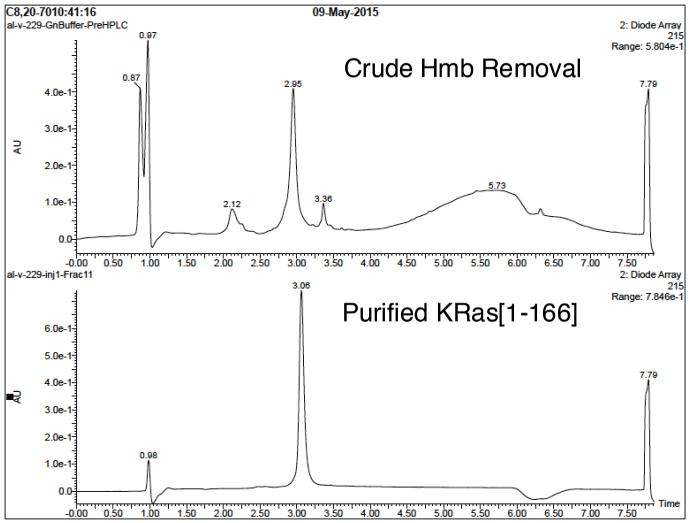


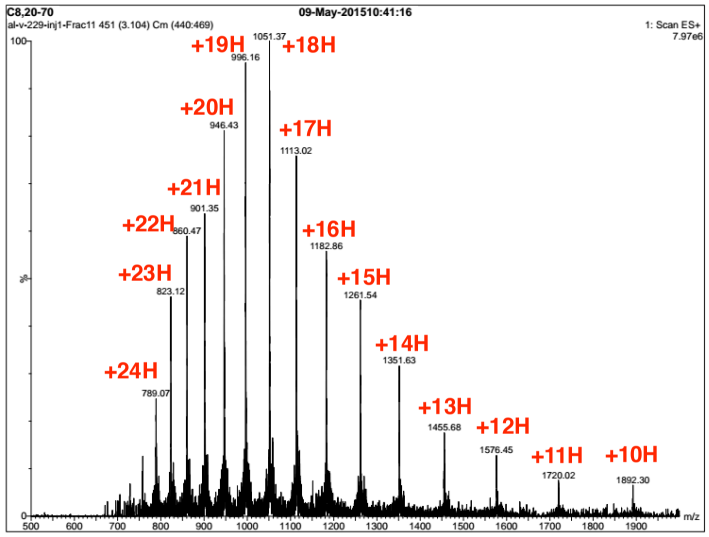
Chemical Formula: $C_{440}H_{702}N_{128}O_{135}S_3$
Molecular Weight: 10041.40



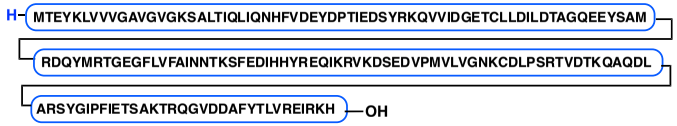
KRas [80-166] (**18**)

UPLC/MS and ESI(+)-MS Spectra of **13**



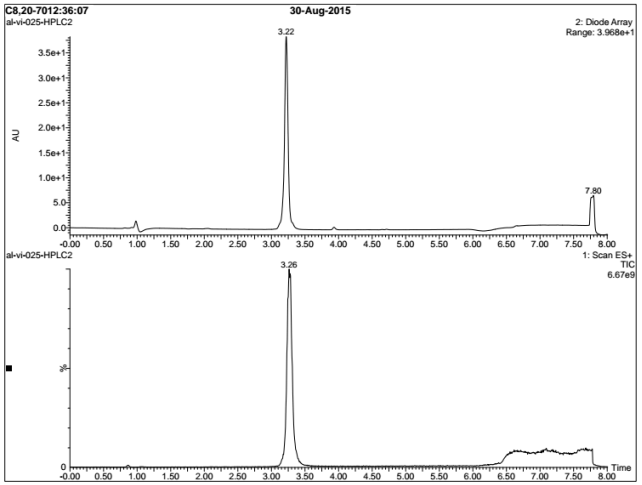


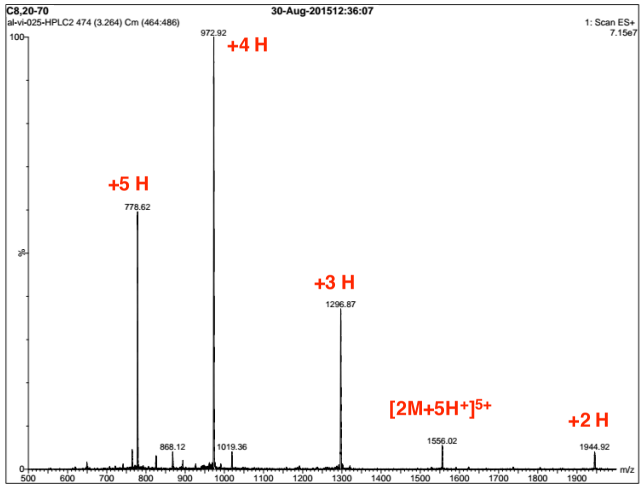
Chemical Formula: $C_{831}H_{1315}N_{227}O_{262}S_7$
Molecular Weight: 18902.41



KRas [1-166] (**13**)

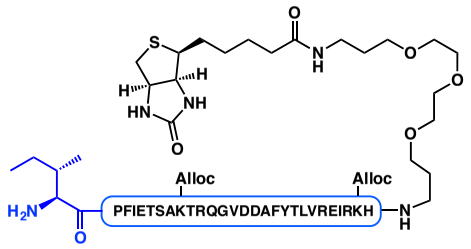
UPLC/MS and ESI(+)-MS Spectra of **19**





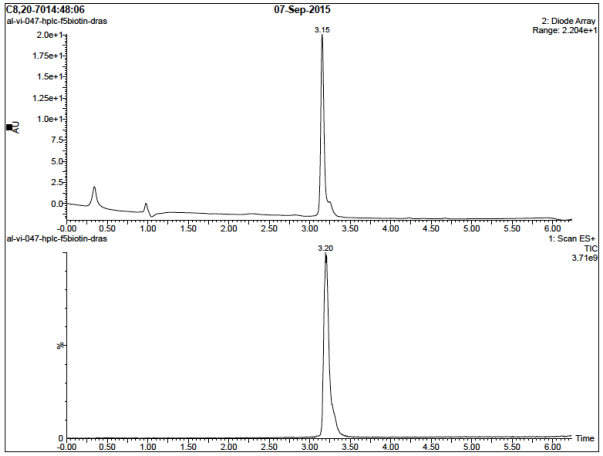
Chemical Formula: $C_{176}H_{280}N_{46}O_{51}S$
Molecular Weight: 3888.51

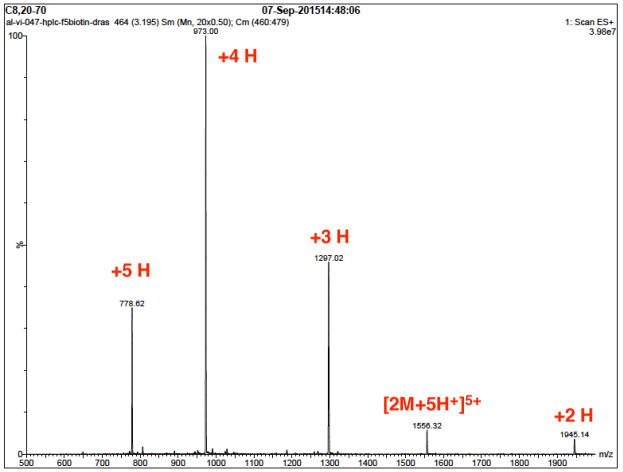
(All-L residues)



KRas[139-166] (**19**)

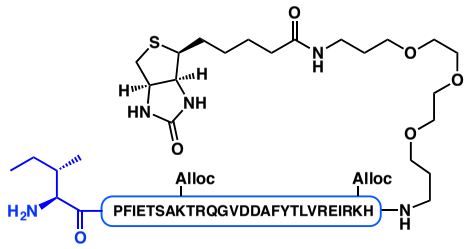
UPLC/MS and ESI(+)-MS Spectra of **19**





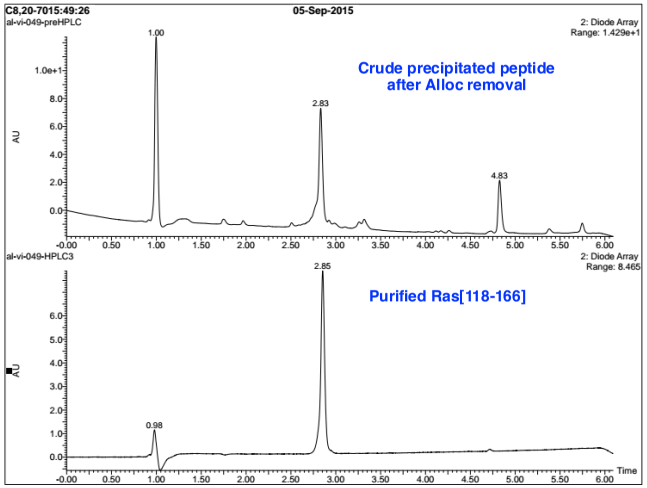
Chemical Formula: $C_{176}H_{280}N_{46}O_{51}S$
Molecular Weight: 3888.51

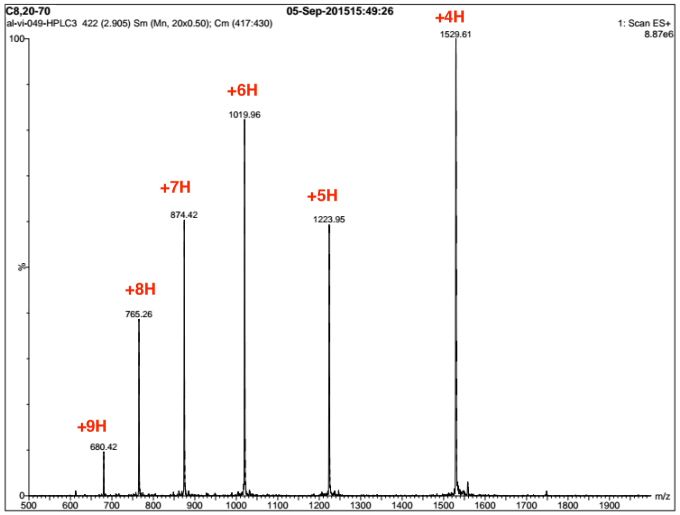
(All-D residues)



KRas[139-166] (**19**)

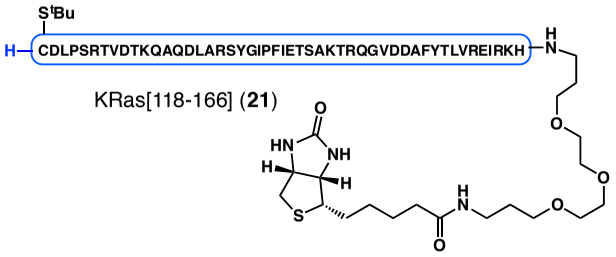
UPLC/MS and ESI(+)-MS Spectra of **21**



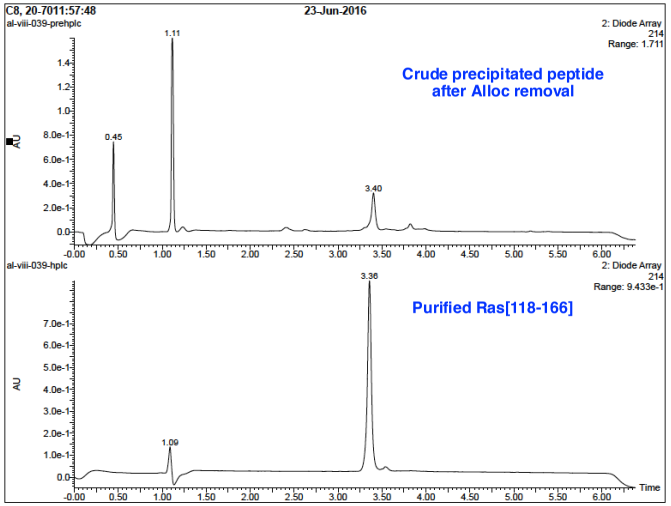


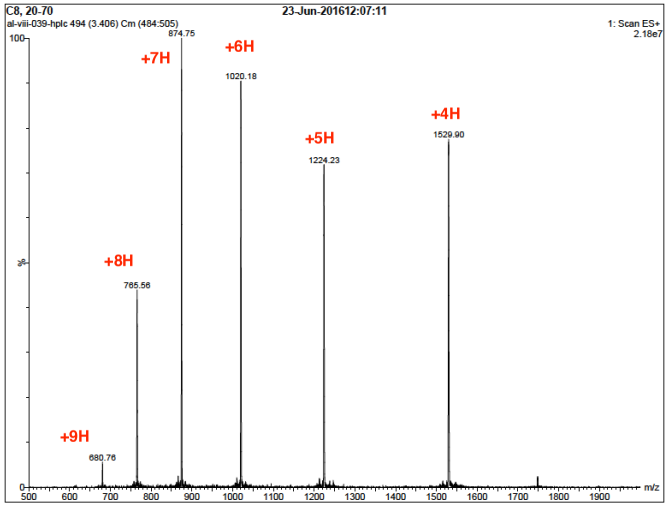
Chemical Formula: $C_{268}H_{436}N_{76}O_{81}S_3$
Molecular Weight: 6115.07

(All-L residues)



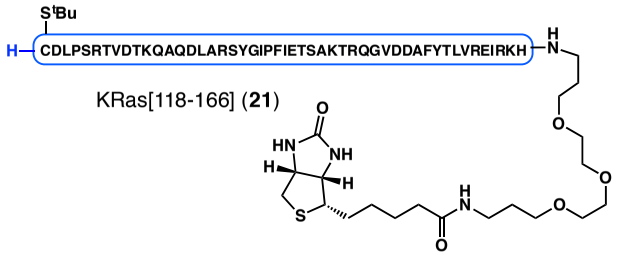
UPLC/MS and ESI(+)-MS Spectra of **21**



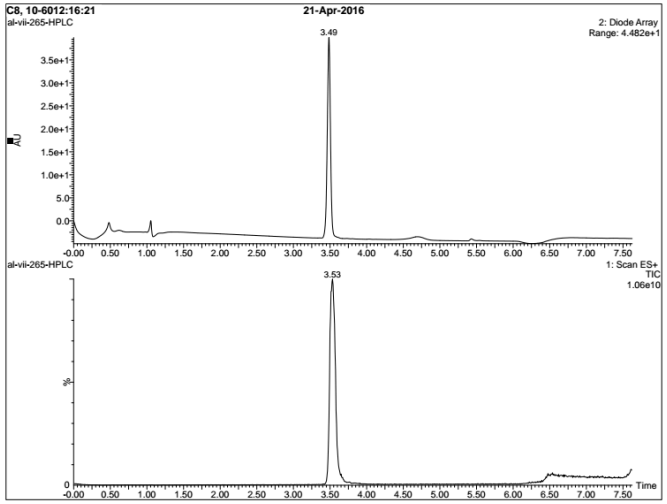


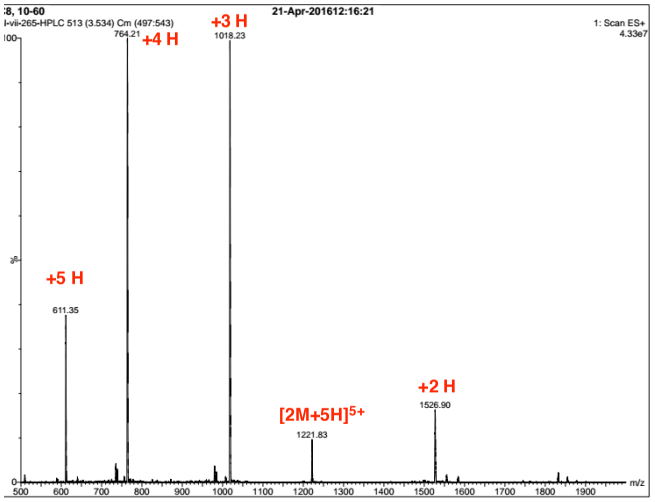
Chemical Formula: $C_{268}H_{436}N_{76}O_{81}S_3$
Molecular Weight: 6115.07

(All-D residues)

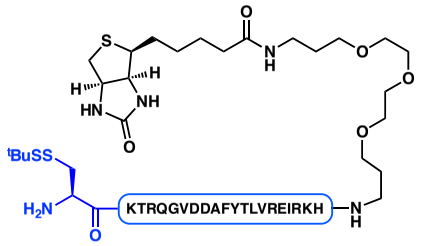


UPLC/MS and ESI(+)-MS Spectra of **22**



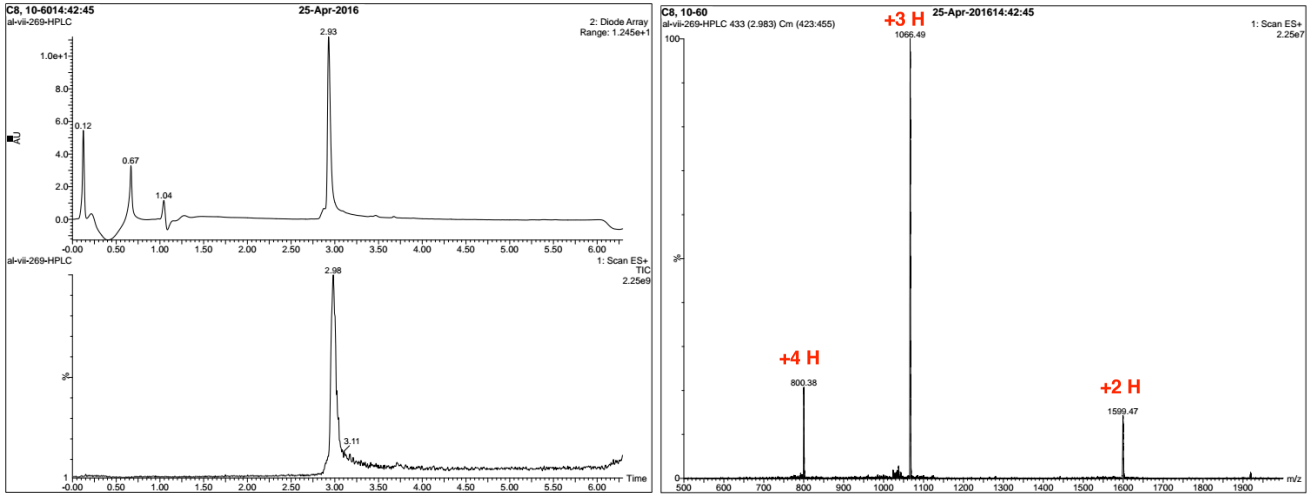


Chemical Formula: $C_{134}H_{223}N_{39}O_{36}S_3$
Molecular Weight: 3052.68

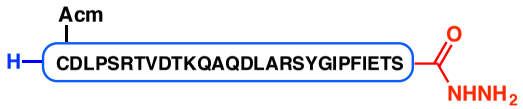


KRas[146-166] (**22**)

UPLC/MS and ESI(+)-MS Spectra of **S-1**

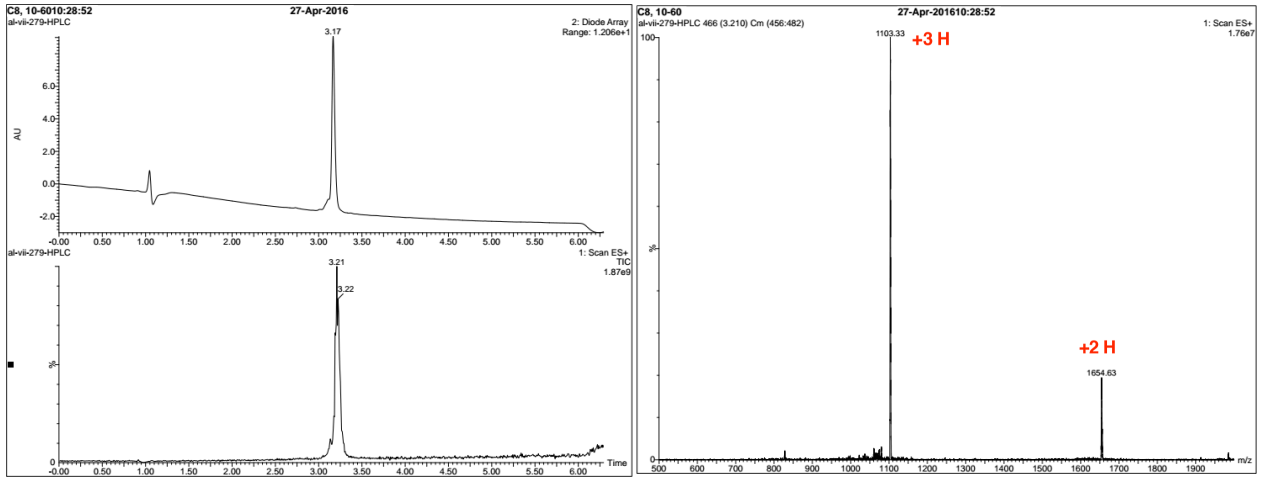


Chemical Formula: $C_{137}H_{222}N_{40}O_{46}S$
Molecular Weight: 3197.58



KRas[118-145] (**S-1**)

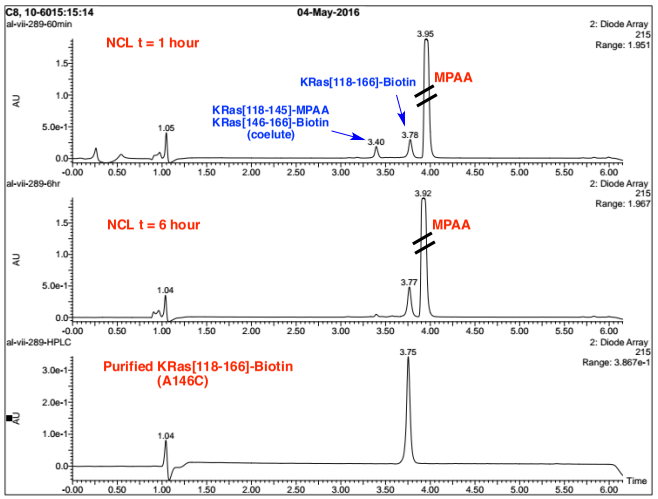
UPLC/MS and ESI(+)-MS Spectra of **23**

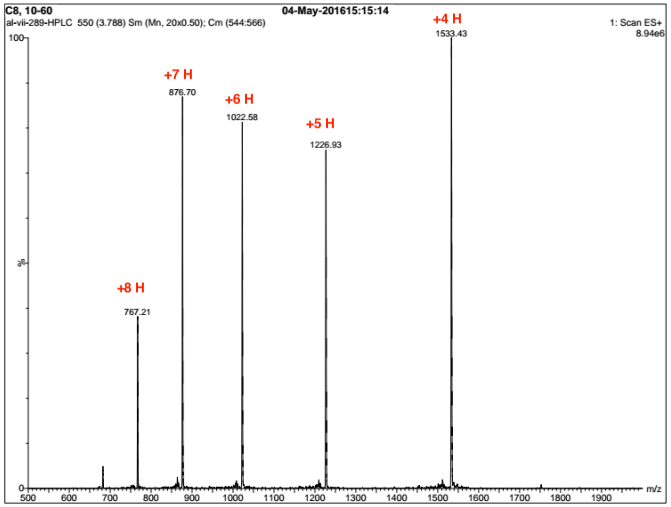


Chemical Formula: $C_{139}H_{224}N_{38}O_{49}S_3$
Molecular Weight: 3307.72

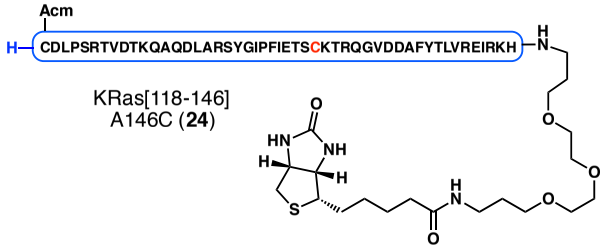
KRas[118-145] (**23**)

UPLC/MS and ESI(+)-MS Spectra of **24**

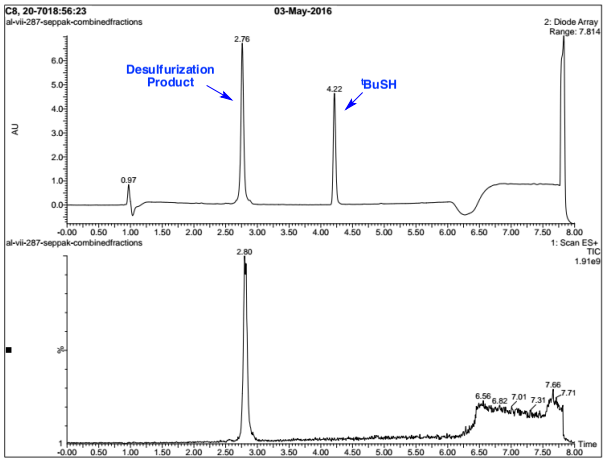


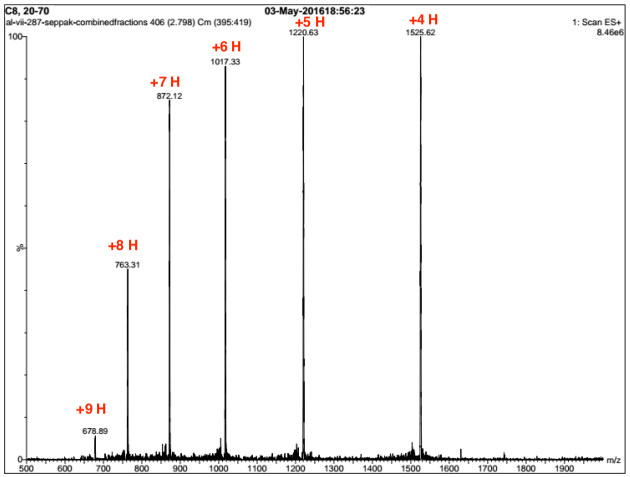


Chemical Formula: $C_{267}H_{433}N_{77}O_{82}S_3$
Molecular Weight: 6130.04

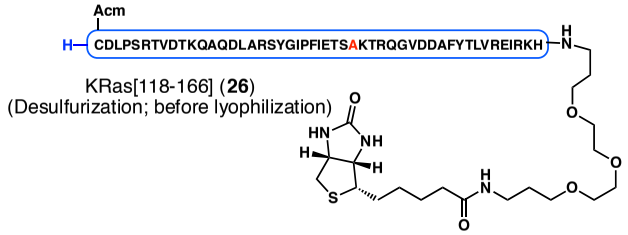


UPLC/MS and ESI(+)-MS Spectra of **24 (desulfurized)**

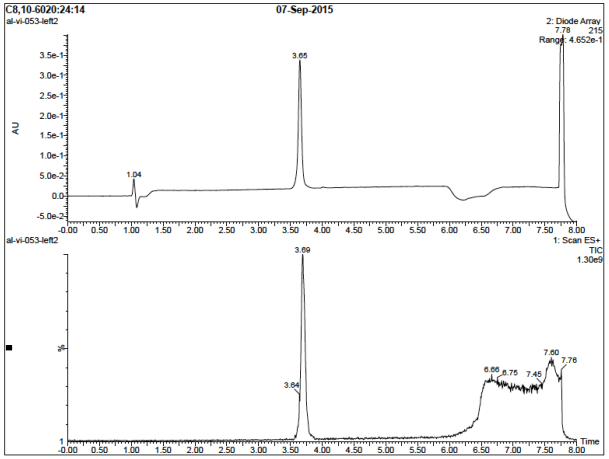


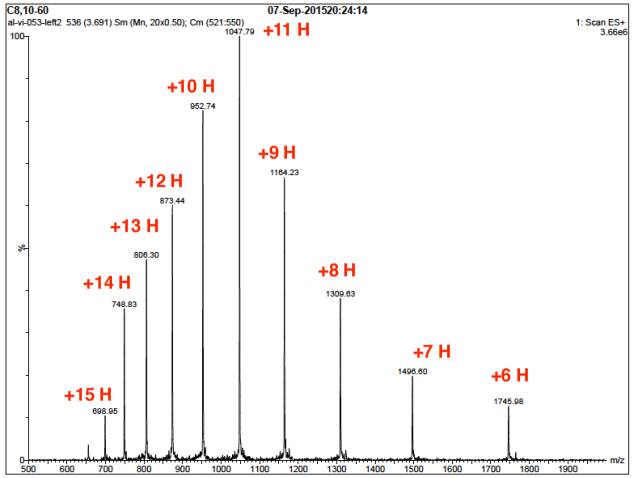


Chemical Formula: $C_{267}H_{433}N_{77}O_{82}S_2$
Molecular Weight: 6097.98



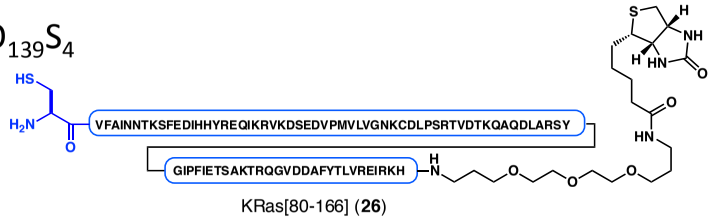
UPLC/MS and ESI(+)-MS Spectra of **26**





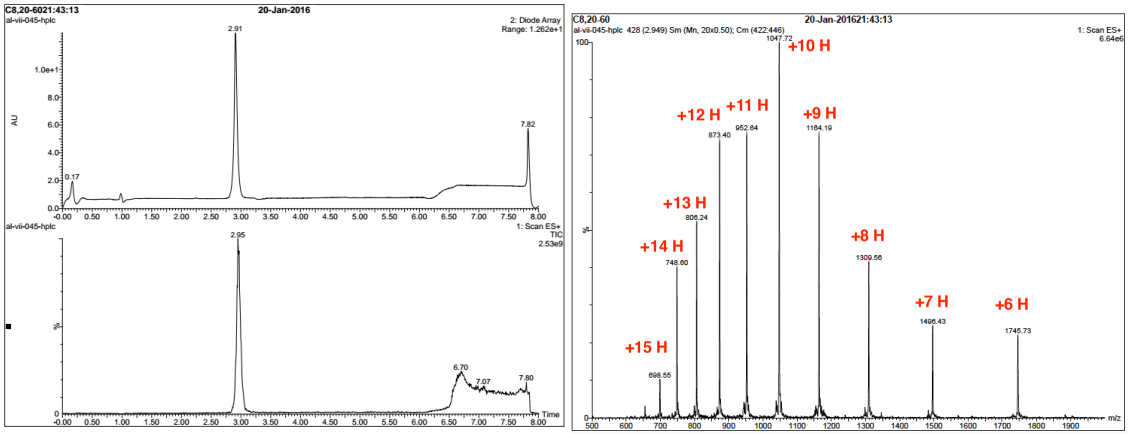
Chemical Formula: $C_{460}H_{738}N_{132}O_{139}S_4$
Molecular Weight: 10469.99

(all-L residues)



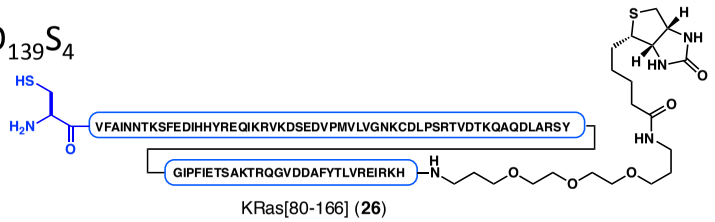
KRas[80-166] (**26**)

UPLC/MS and ESI(+)-MS Spectra of **26**



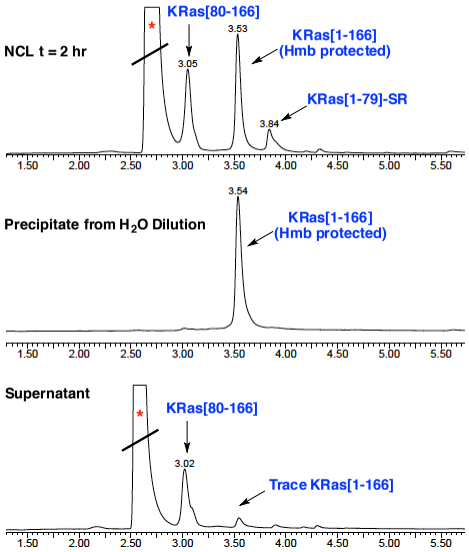
Chemical Formula: $C_{460}H_{738}N_{132}O_{139}S_4$
Molecular Weight: 10469.99

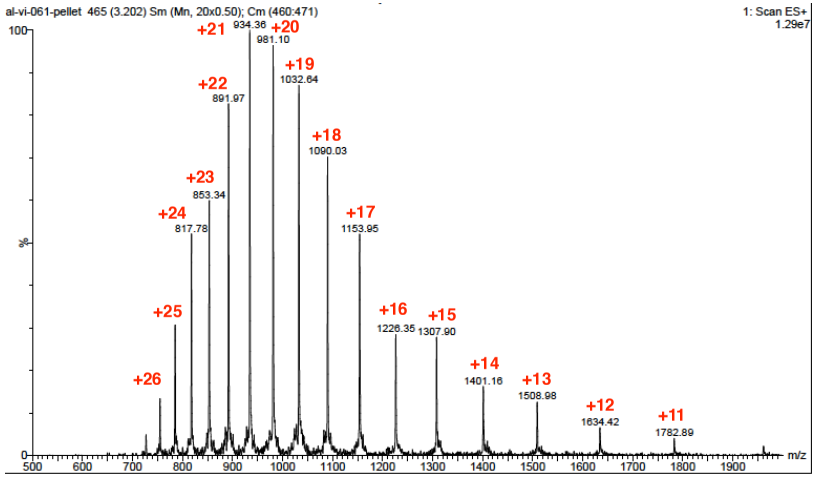
(All-D Residues)



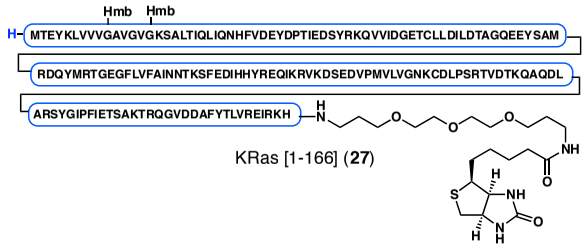
KRas[80-166] (**26**)

UPLC/MS and ESI(+)-MS Spectra of **27**

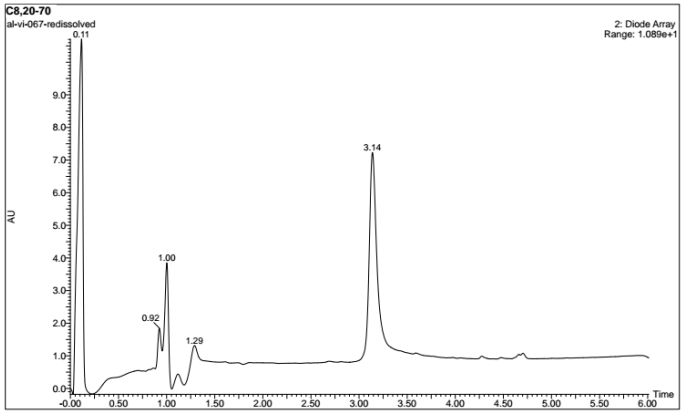


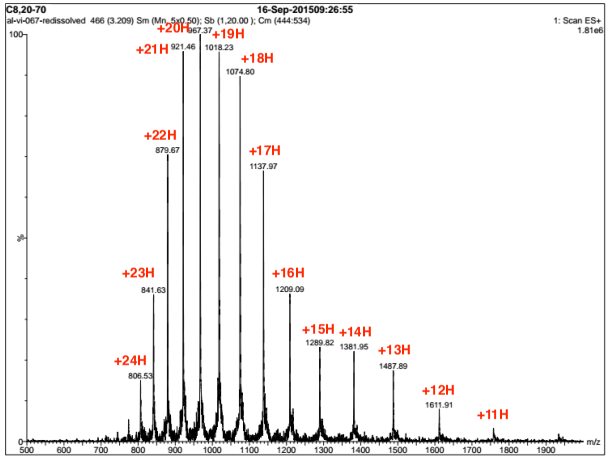


Chemical Formula: $C_{867}H_{1367}N_{231}O_{270}S_8$
Molecular Weight: 19603.30



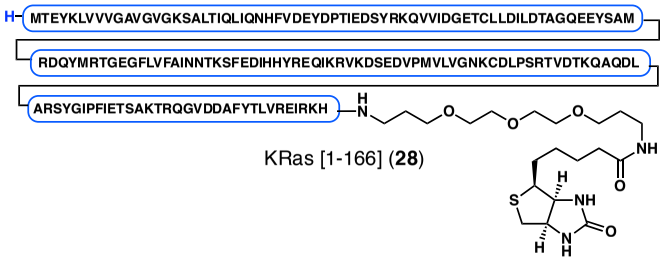
UPLC/MS and ESI(+)-MS Spectra of **28** (all L-residues)





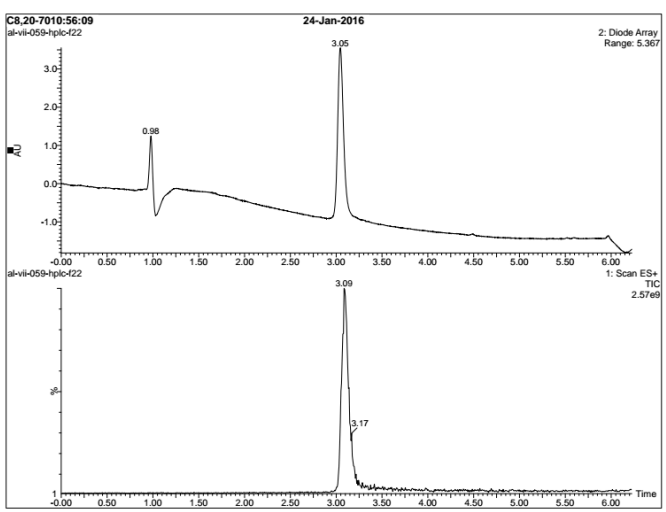
Chemical Formula: $C_{851}H_{1351}N_{231}O_{266}S_8$
Molecular Weight: 19331.00

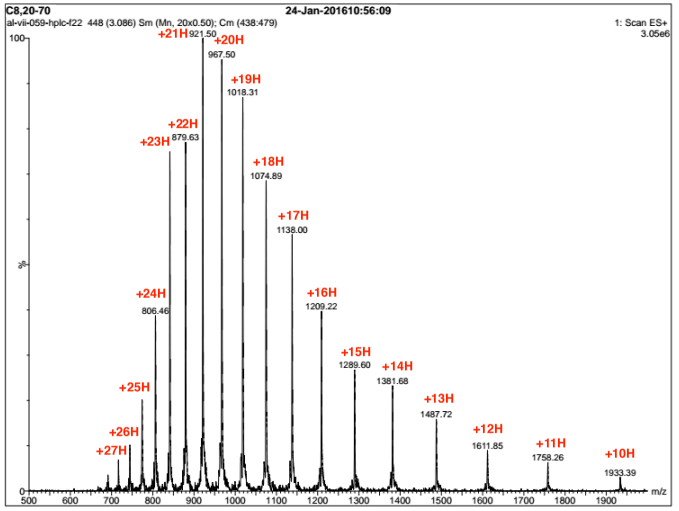
(all-L residues)



KRas [1-166] (**28**)

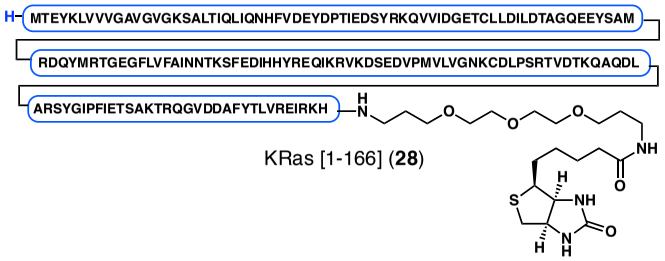
UPLC/MS and ESI(+)-MS Spectra of **28** (all D-residues)



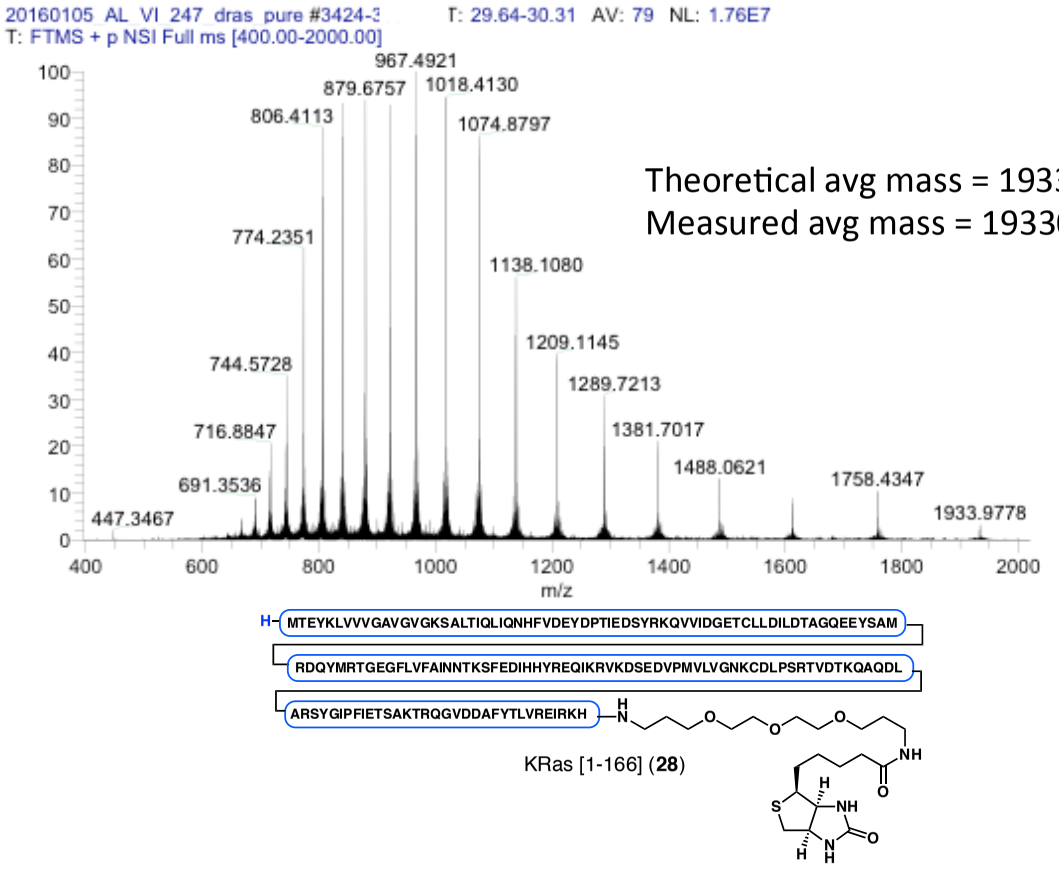


Chemical Formula: $C_{851}H_{1351}N_{231}O_{266}S_8$
Molecular Weight: 19331.00

(all-D residues)

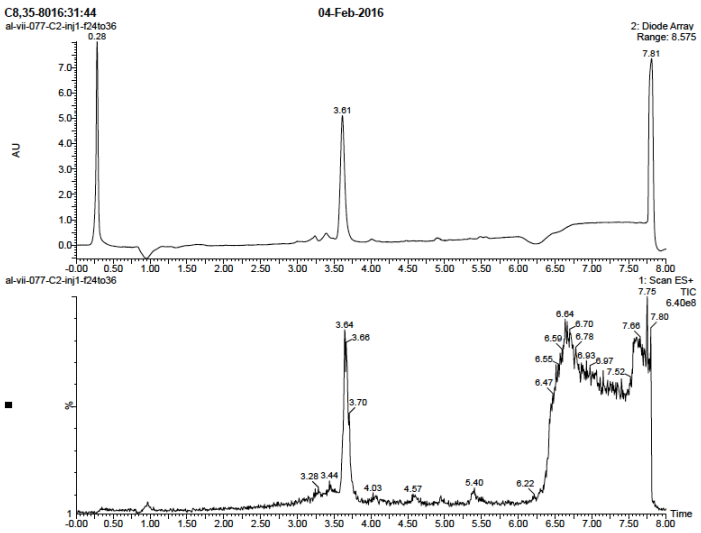


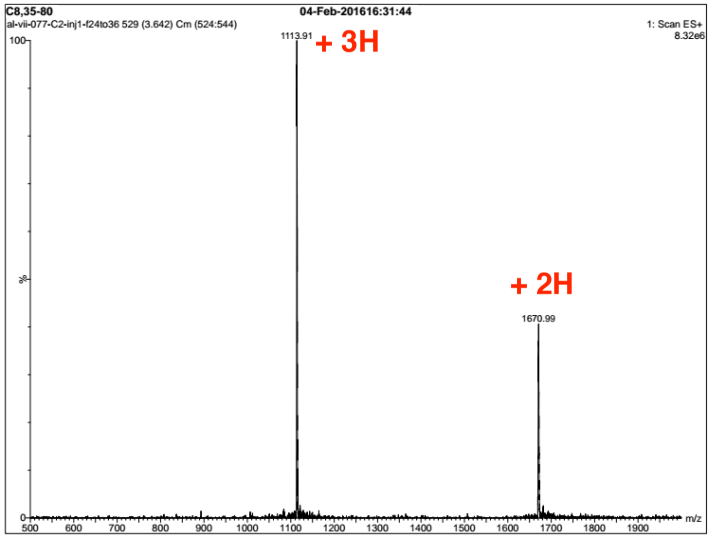
High Resolution MS of **28** (all D-residues)



Theoretical avg mass = 19330.74
Measured avg mass = 19330.4233

UPLC/MS and ESI(+)-MS Spectra of **42**

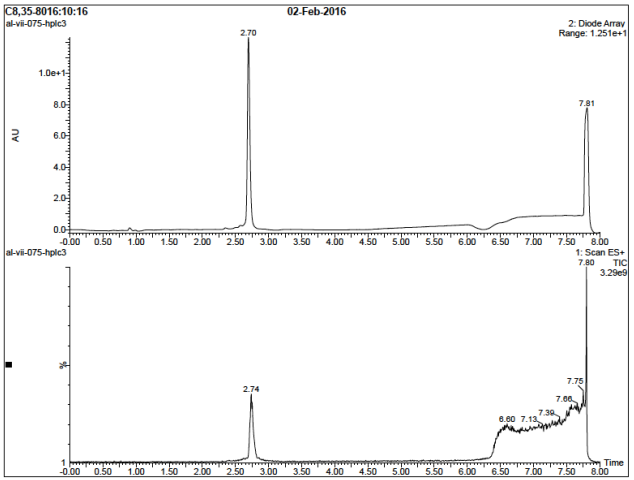


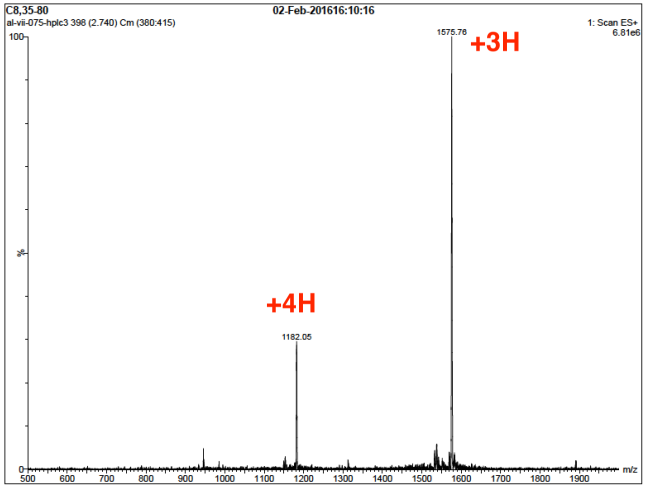


Chemical Formula: $C_{156}H_{244}N_{38}O_{41}S$
MW = 3339.95 g/mol

HIV[1-27] (**42**)

UPLC/MS and ESI(+)-MS Spectra of **43**

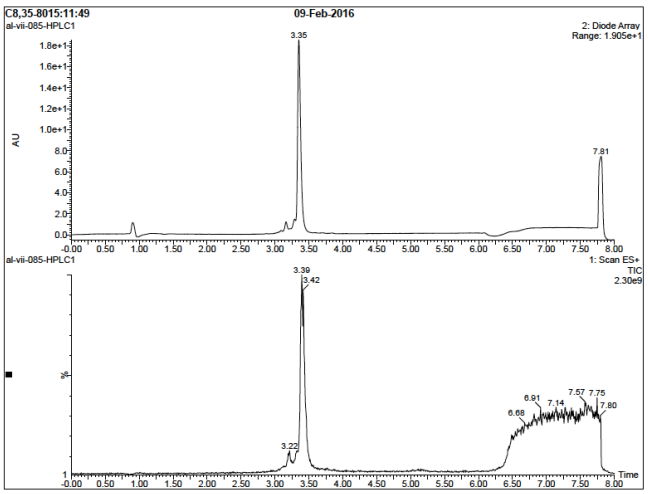


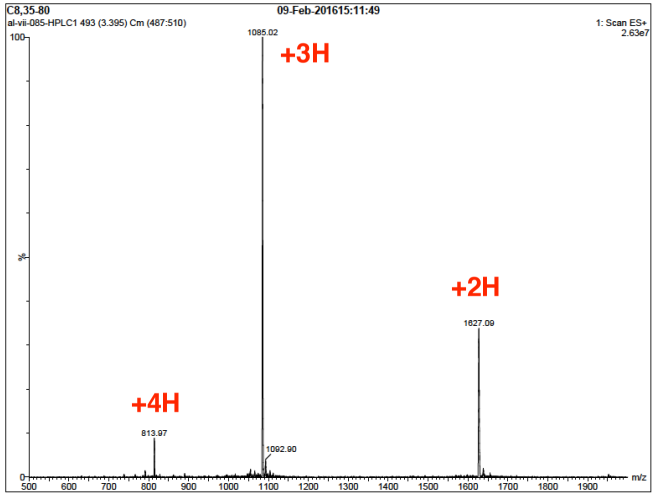


Chemical Formula: $C_{214}H_{332}N_{50}O_{64}S_3$
MW = 4725.48 g/mol

HIV[28-66] (**43**)

UPLC/MS and ESI(+)-MS Spectra of **44**

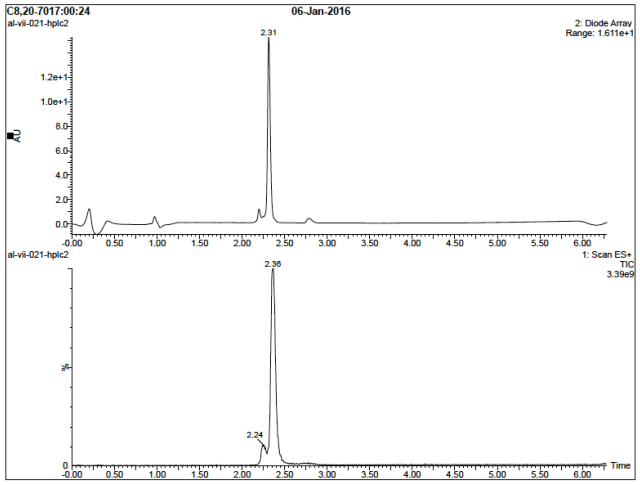


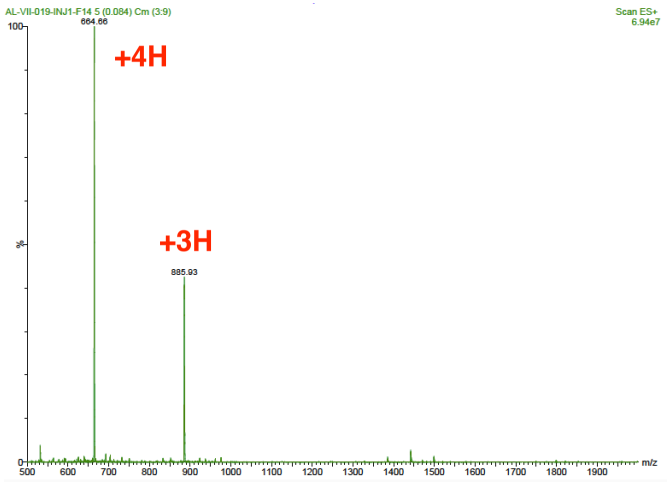


Chemical Formula: $C_{149}H_{239}N_{37}O_{38}S_3$
MW = 3252.95 g/mol

HIV[67-94] (**44**)

UPLC/MS and ESI(+)-MS Spectra of **45**

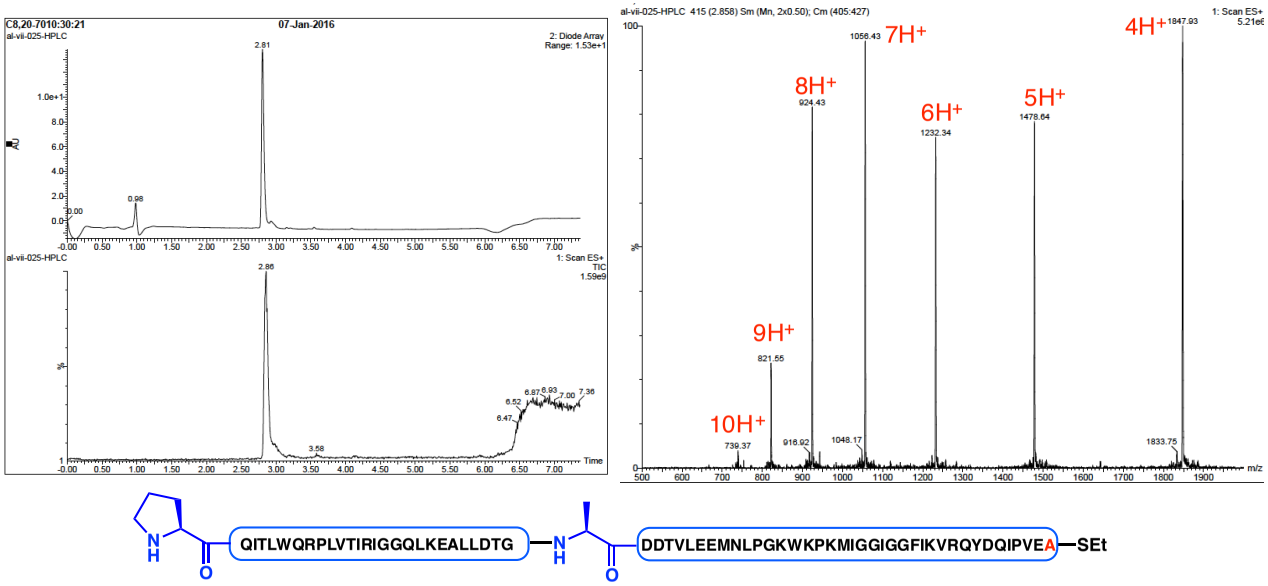




Chemical Formula: $C_{115}H_{201}N_{41}O_{27}S_2$
MW = 2654.25 g/mol

HIV[95-115] (**45**)

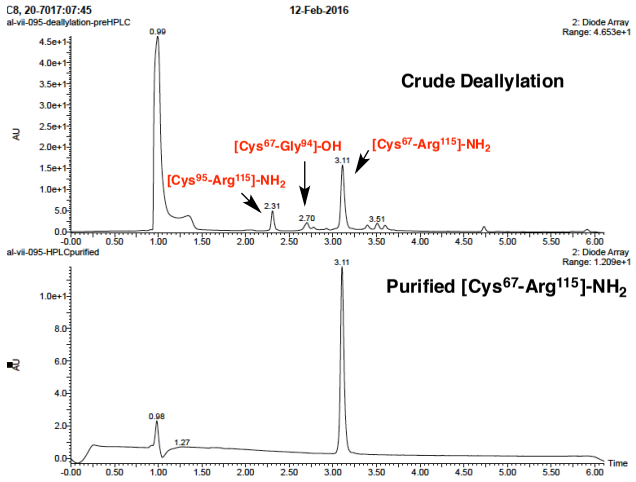
UPLC/MS and ESI(+)-MS Spectra of **46**

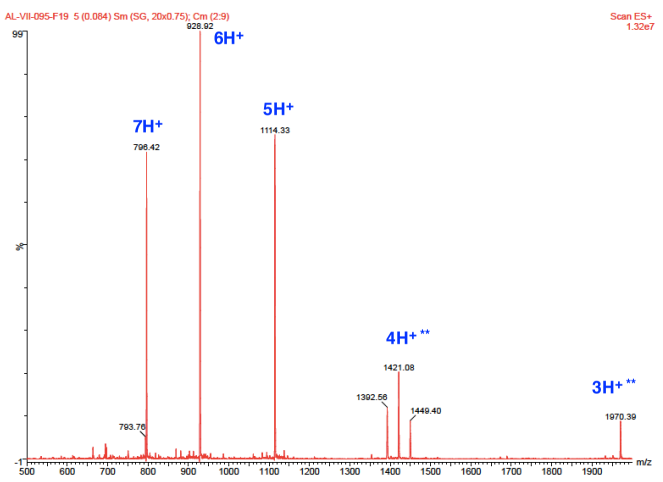


HIV-1 Protease I66A [1-66] (**46**)

Chemical Formula: $C_{335}H_{544}N_{88}O_{93}S_3$
MW = 7388.74 g/mol

UPLC/MS and ESI(+)-MS Spectra of **47**



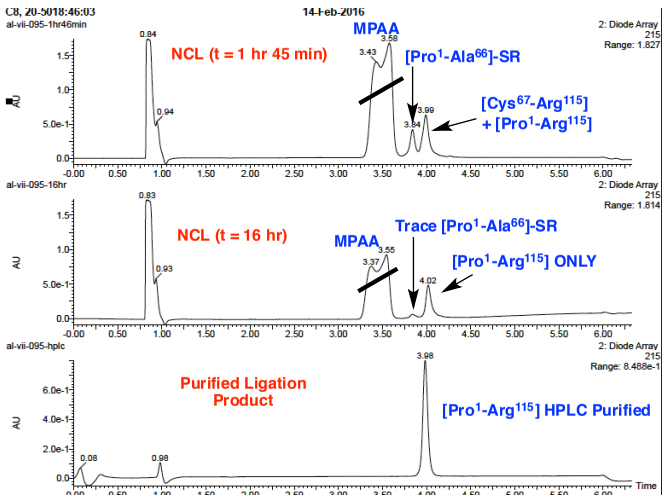


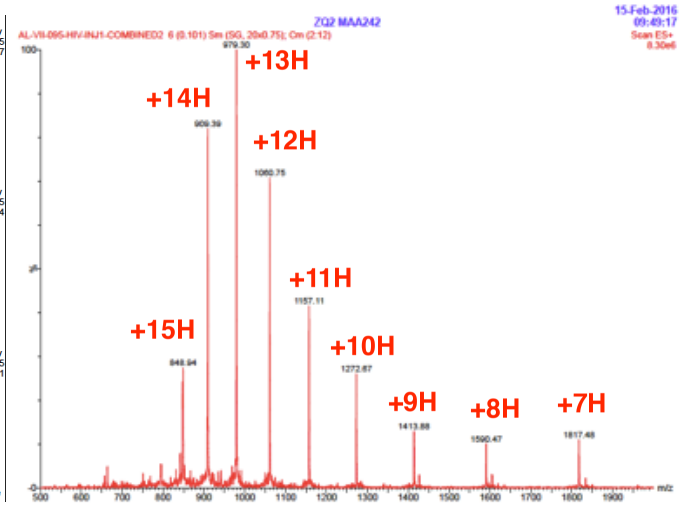
** Multiple peaks due to association of 1-3 TFA molecules

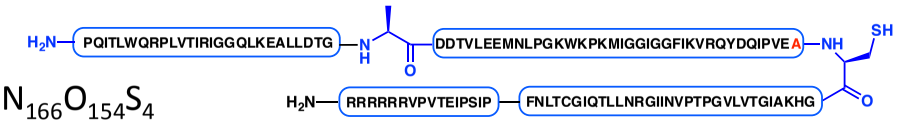
Chemical Formula: $C_{245}H_{424}N_{78}O_{61}S_4$
MW = 5566.81 g/mol

HIV[67-115] (**47**)

UPLC/MS and ESI(+)-MS Spectra of **48**



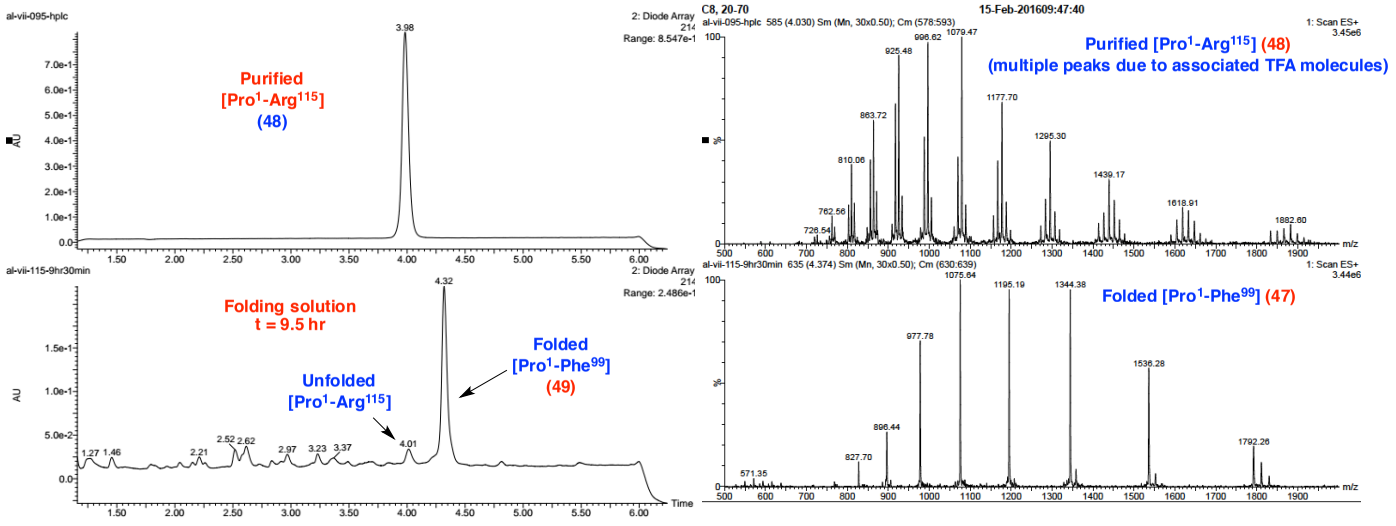


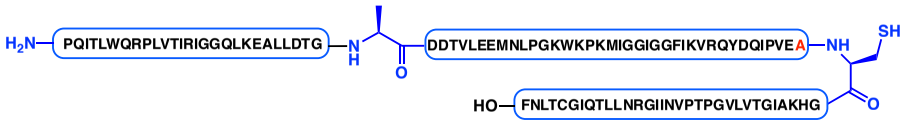


HIV[1-115] (**48**)

Chemical Formula: $C_{570}H_{946}N_{166}O_{154}S_4$
MW = 12707.09 g/mol

UPLC/MS and ESI(+)-MS Spectra of **49**





HIV[1-99] (**49**)

Chemical Formula: $C_{570}H_{946}N_{166}O_{154}S_4$
MW = 12707.09 g/mol